520 I52p2 68-54331
Inglis
Planets, stars, and galaxies

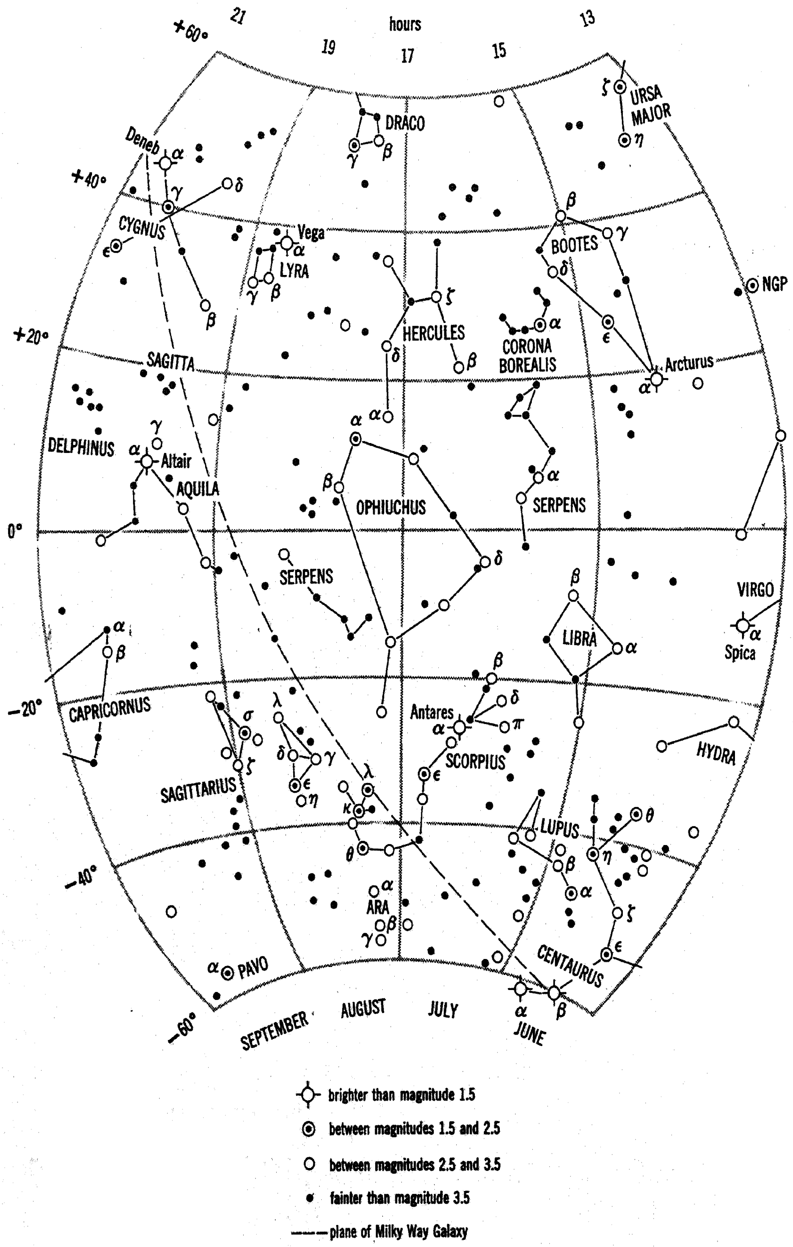
MAIN

kansas city public library
kansas city, missouri

Books will be issued only
on presentation of library card.
Please report lost cards and
change of residence promptly.
Card holders are responsible for
all books, records, films, pictures
or other library materials
checked out on their cards.

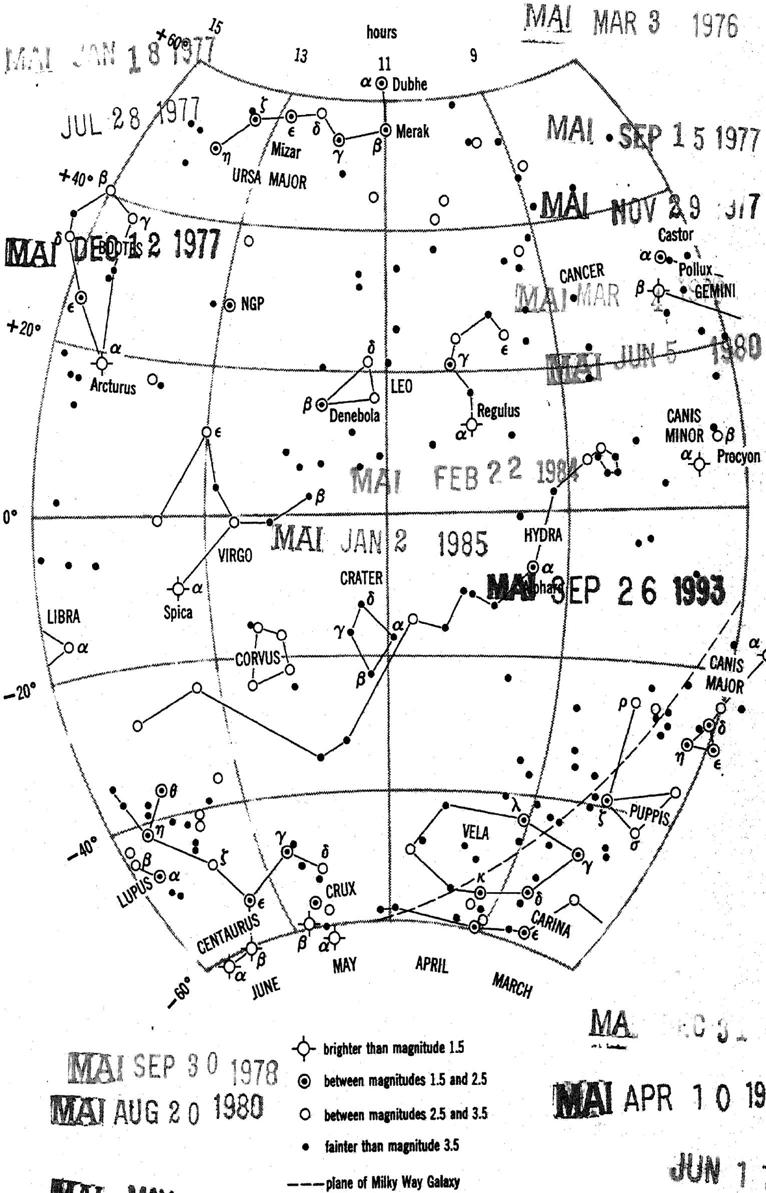
MAIN

THE SUMMER CONSTELLATIONS



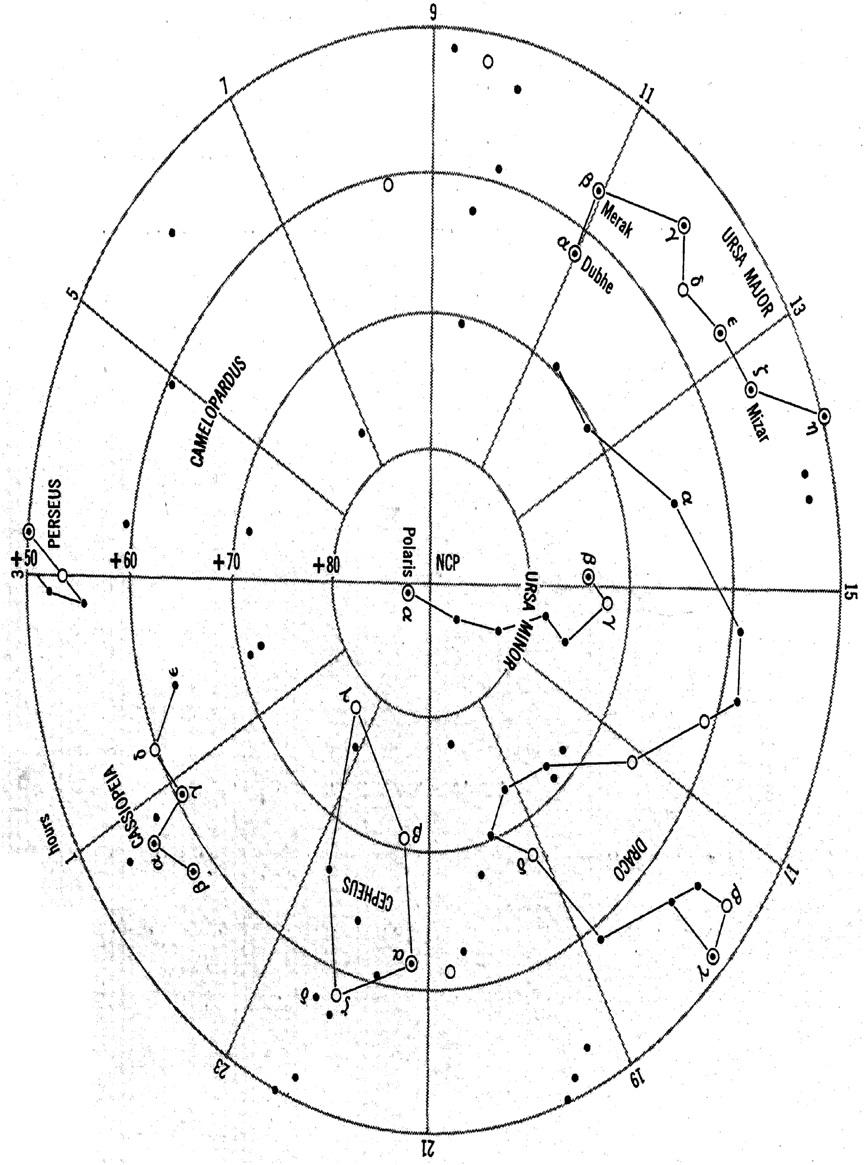
MAI JAN 27 1976
3 1148 00099 1851

THE SPRING CONSTELLATIONS



MAI MAR 3 1976
MAI JAN 18 1977
JUL 28 1977
MAI SEP 15 1977
MAI NOV 29 1977
MAI DEC 12 1977
MAI MAR 2
MAI JUN 5 1980
MAI FEB 22 1984
MAI JAN 2 1985
MAI SEP 26 1993
MA 7
MAI SEP 30 1978
MAI AUG 20 1980
MAI APR 10 1993
MAI MAY 17 1986
JUN 17 199

THE NORTHERN CONSTELLATIONS



-○- brighter than magnitude 1.5

⊙ between magnitudes 1.5 and 2.5

○ between magnitudes 2.5 and 3.5

● fainter than magnitude 3.5

DATE DUE
OCT 14 1975
MAI NOV 8 1976
MAI SEP 15 1977
MAI AUG 5 1978
MAI NOV 10 [illegible]

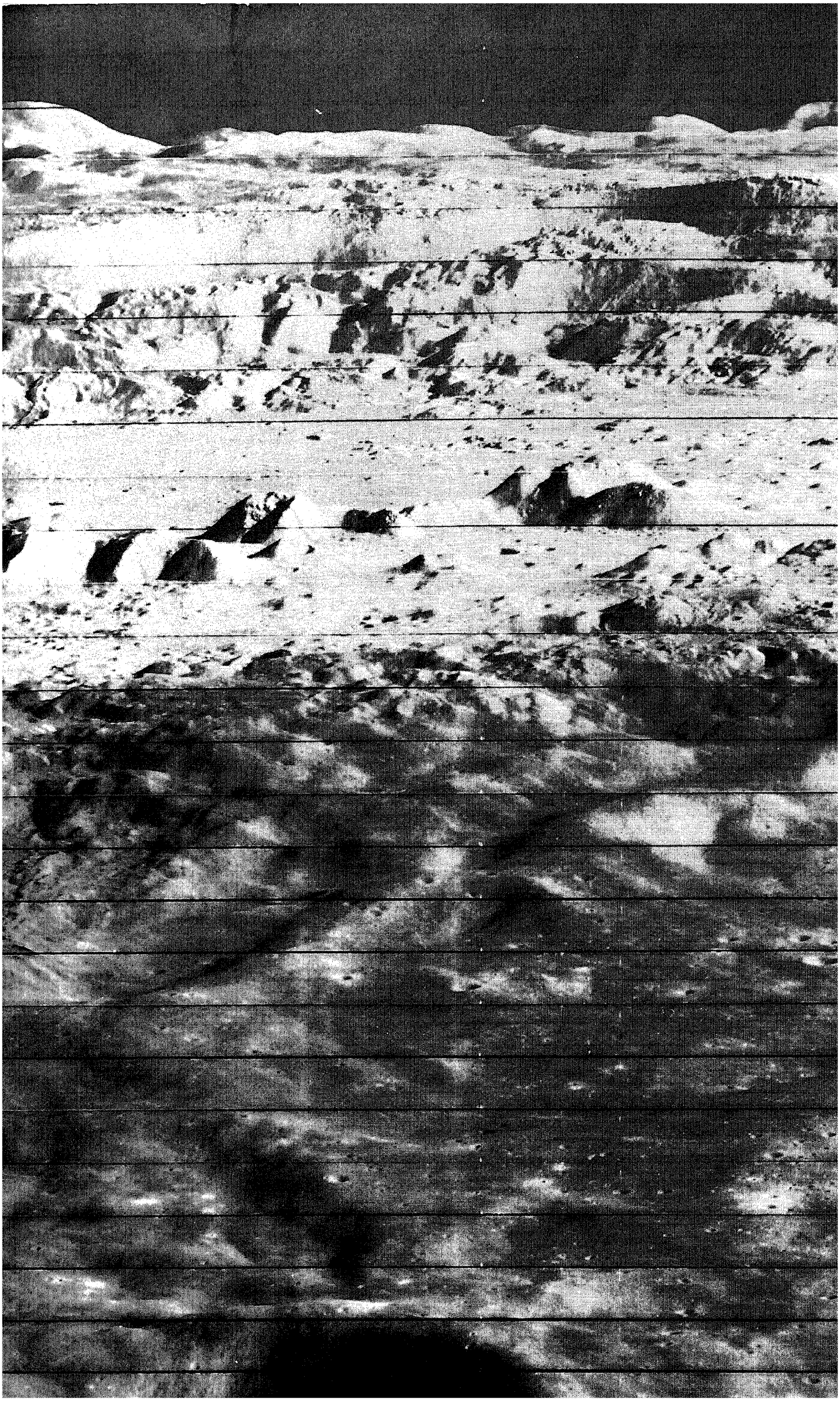
On the following pages:

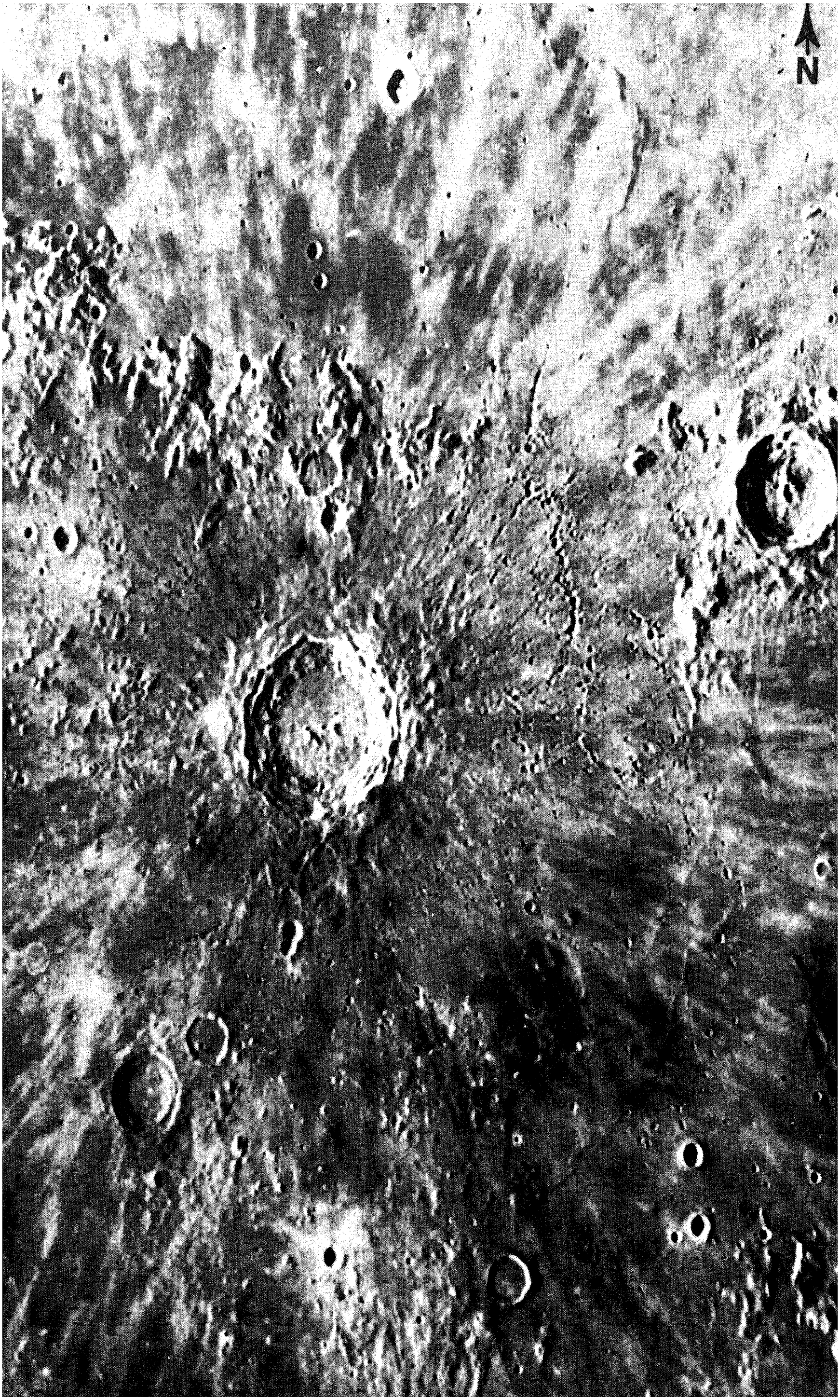
First. A portion of the crater Copernicus as photographed by Lunar Orbiter II on November 23, 1966. In the foreground can be seen the double-crater Fauth, 13 miles long and 4,500 ft deep. Beyond Fauth the landscape slopes up to the rim of Copernicus (about the middle of the photograph). Beyond the rim is the crater floor with the central peaks projecting out of it. Beyond the peaks is a smoother portion of the floor and then the wall of the crater itself with steep cliffs over 1,000 ft high and angular blocks 150 ft across. On the horizon more distant mountains can be seen.

It is about 180 miles from the bottom to the top of the photograph and about 17 miles across near the middle. The crater Copernicus is 60 miles in diameter from rim to rim and about 2 miles deep. (NASA photograph)

Second. The crater Copernicus photographed from the Lick Observatory. (Lick Observatory photograph)

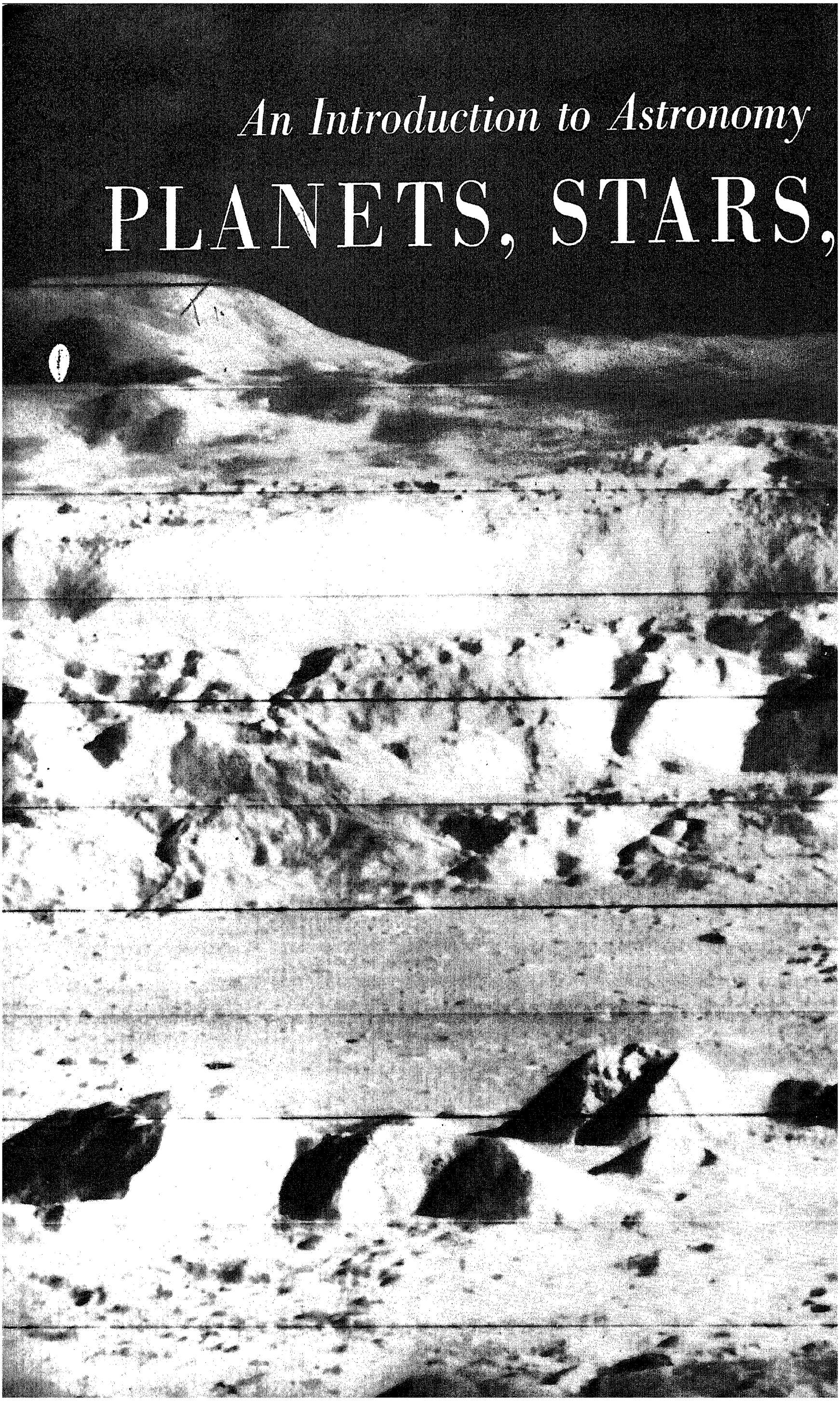
Third. An enlargement of the photograph of the crater Copernicus taken by Lunar Orbiter II. The spacecraft was 28.4 miles above the surface of the moon and 150 miles from the crater when the photograph was taken.




N

An Introduction to Astronomy

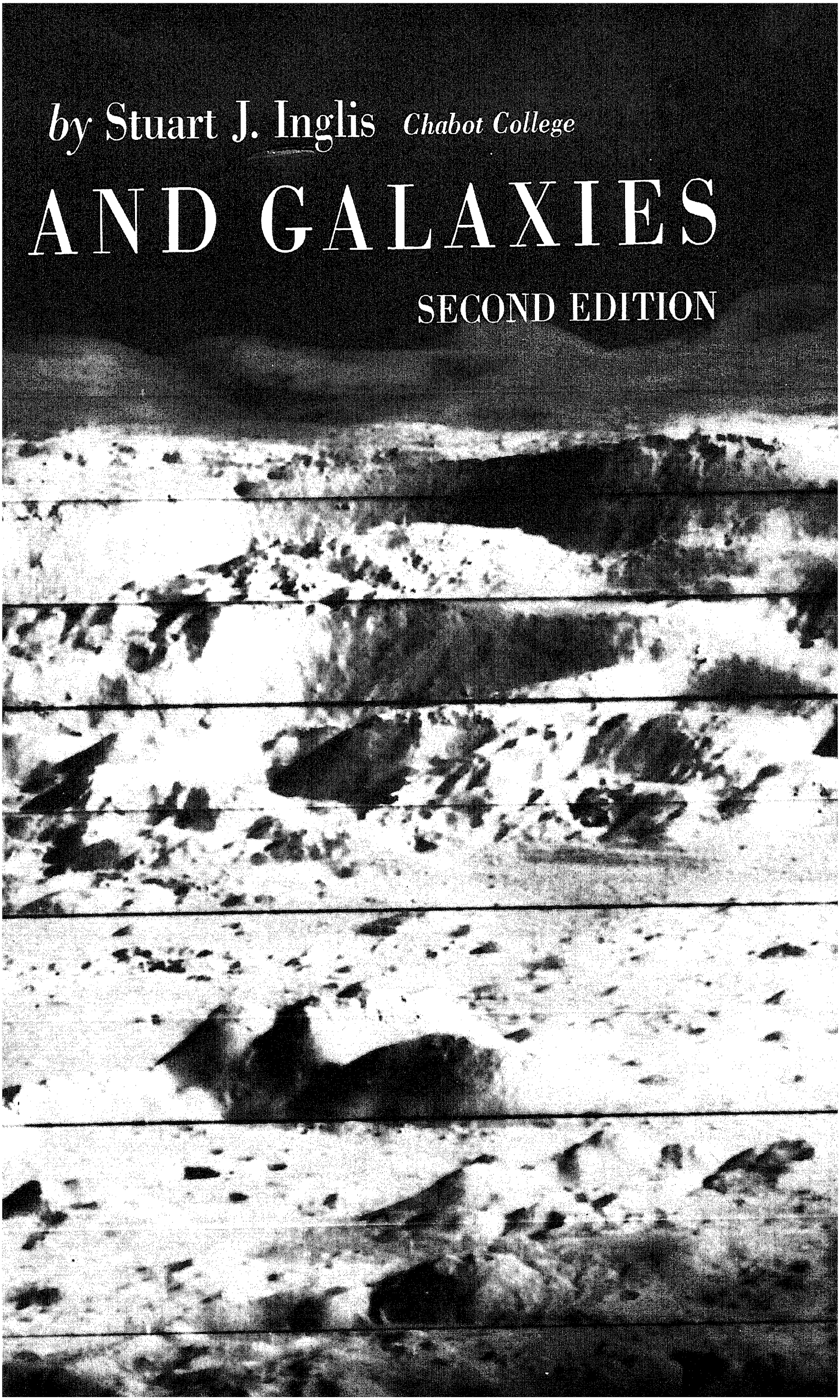
PLANETS, STARS,



by Stuart J. Inglis *Chabot College*

AND GALAXIES

SECOND EDITION



FIRST CORRECTED PRINTING, AUGUST, 1967

Copyright © 1961, 1967 *by John Wiley & Sons, Inc.* All Rights Reserved. This book or any part thereof must not be reproduced in any form without the written permission of the publisher. Library of Congress Catalog Card Number: 67-17342. Printed in the United States of America

To ISA

KANSAS CITY (MO.) PUBLIC LIBRARY

6854331

PREFACE

The purpose of this book has not changed in the new edition: there is still a need to present the subject of astronomy to those who, without being well-grounded in mathematics, nevertheless feel that man must have some sense of his place in and his relation to the physical universe. With the ever-increasing ability of man to observe and to travel in the universe it becomes even more important that knowledge of astronomy not be restricted to the scientist.

I have revised this book with continued interest in the liberal arts student. It is intended to be used as a college text for a one-semester or one-quarter course. It is my experience that the liberal arts student today is better prepared than students of ten or even five years ago. He is better able to grasp concepts and in fact he is more apt to insist that he understand. Because of this, I have eliminated some of the less rigorous parts and introduced some new observations and ideas which demand more of the student.

The chapters devoted to the solar system have been extensively re-written. New techniques and new observations have stimulated study of the solar system. To live in the years of the inauguration of interplanetary space probes is exciting; arousing student interest is hardly a problem to the teacher. This increased activity in astronomy and space science, however, makes it more difficult for the student to learn solely from the textbook, for textbooks cannot be entirely up-to-date. New advances are apt to occur even while the book is in press.

The student must rely upon the teacher who keeps abreast of the latest developments. The student may also read from periodicals.

It is interest in the latest developments in astronomy which has encouraged me to let the student become more involved in the subject. I feel that he can greatly benefit by writing a short paper on a subject of his interest. With this in mind I have greatly expanded the bibliography at the end of each chapter. These bibliographies are prepared for the student's use and so contain repeated references to *Sky and Telescope* and *Scientific American.*

I have introduced some mathematics into several sections: in the development of the concept of energy, in the discussion of Newton's law of gravity, in the discussion of resolving power, and in the Doppler principle. The concept of energy is important for the student's understanding of later aspects of the course. Newton's law of gravity is essential. Radio telescopes and radio interferometers make the subject of resolving power all the more important. Observations of the spectral shift of quasars brings about a need for a deeper understanding of the Doppler principle. In fact, the relativistic Doppler equation is *required* and has been introduced into the chapter on cosmology. In all instances the mathematics is nothing more than simple algebra.

In preparing this revision a number of instructors have made suggestions which have been a great help to me. In particular I would like to thank Otto Barrett, instructor of astronomy and physics at Contra Costa College, San Pablo, California, not only for specific suggestions but also for a decade of stimulating and valuable discussions on astronomy and physics. I would also like to thank Bill Smith, instructor of astronomy at Chabot College for his suggestions. But it is I who accept responsibility for any errors or misstatements which may remain in spite of their comments. Proofreading is a chore and I thank Ulli Staniford for her careful and cheerful help reading the galley proofs.

Perhaps it is only the author who understands the debt of gratitude he owes his family for their help and patience. My son, Jeff has not only helped me directly, but been patient with me. My daughters, Adrienne and Jennifer, have been patient. My wife, Isa, has typed, read, and been very understanding. For all their help and consideration, I am truly grateful.

Stuart J. Inglis

Chabot College
Hayward, California

CONTENTS

1 *Introduction* 3

2 *Basic Tools and Methods* 13

3 *The Atom* 45

4 *The Physical Characteristics of the Planets* 59

5 *The Planetary Motions* 95

6 *Satellites* 117

7 *Minor Planets, Meteorites, and Comets* 149

8 *The Age and Origin of the Solar System* 185

9 *The Sun* 201

10 *The Stars, a General Description* 241

11 *Nonstable Stars* 271

12 *The Evolution of Stars* 303

13 *Multiple Stars and Clusters* 329

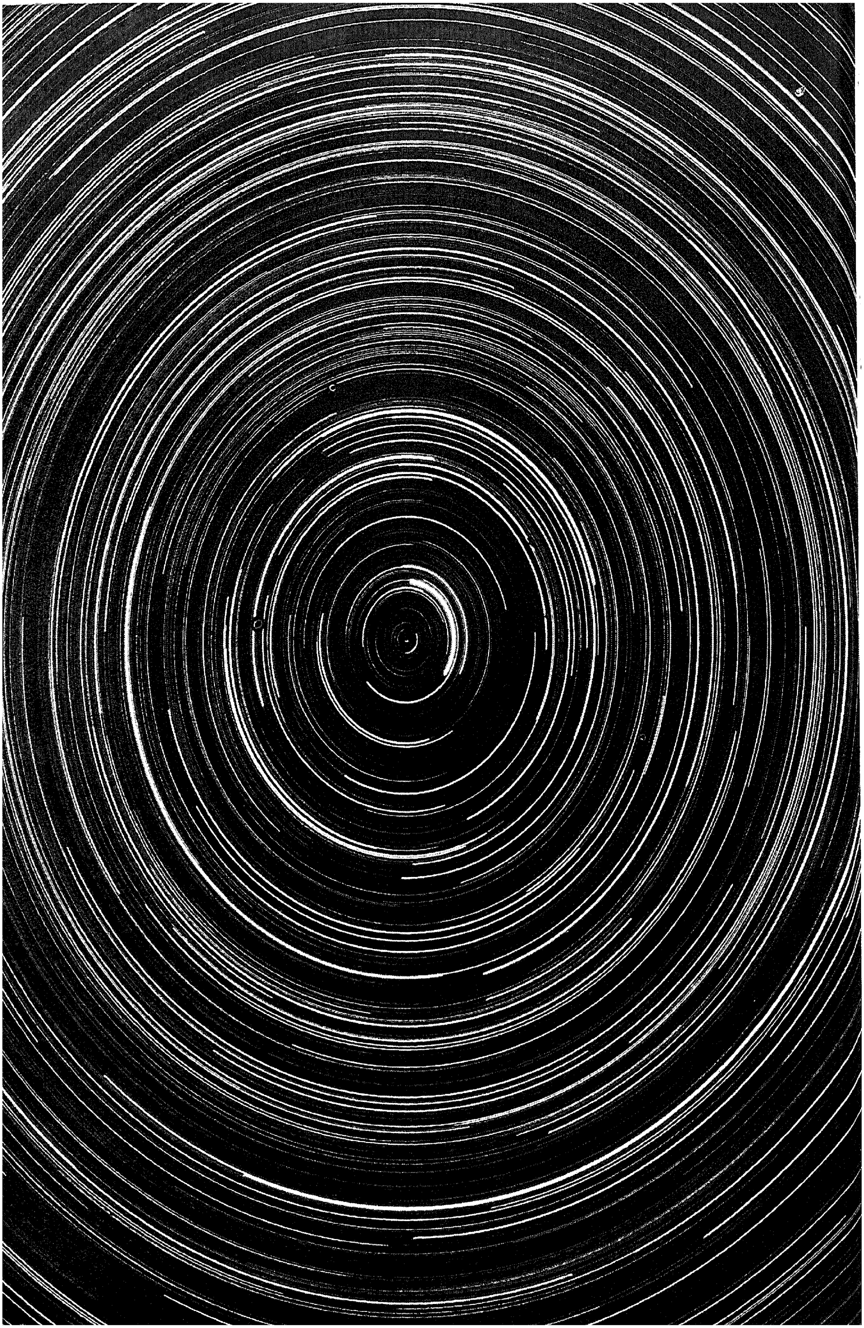
14 *Between the Stars* 367

15 *Galaxies* 389

16 *The Universe and Relativity* 429

Questions and Problems 463

Index 471



Diurnal (daily) motion of the stars about the north celestial pole. This photograph is a time exposure taken with a stationary camera. (Lick Observatory)

Chapter 1 *Introduction*

Every field of study has its own distinguishing features. Each has aspects that are unseen by the average individual and that, because they are unseen, tend to the mysterious. It is this mystery that fascinates us. That which is unseen and not understood presents us with a challenge—the challenge of discovering, of understanding. This is innate in the human mind and has led to the technological developments that identify our civilization. Why else are there miracle drugs to heal us, jet planes to carry us through the air?

Most fields of study are influenced by our desire to create either a better or a more convenient life. Among these are medicine, engineering, psychology, meteorology, and political science.

There remain a few scientific disciplines, however, that are free from the restriction of trying to improve our everyday existence. These are called pure sciences. A pure science may result in some understanding that will ultimately help mankind, but the basic *motive* for wanting to understand is pure curiosity.

Astronomy is such a science. To study the distant stars and galaxies has nothing whatever to do with making our daily life more comfortable and convenient. In a sense, it is completely impractical. Yet astronomy has influenced our lives tremendously. We can see the best example of this influence in the work of Sir Isaac Newton.

Newton, an astronomer and mathematician of the seventeenth century, sought to explain the fact that the moon circles the Earth and does not fly off into space. He approached the problem by making a mathematical analysis of the moon's motion. The result of this study was his laws of motion and the famed law of gravitation, by which we can predict not only the motion of the moon but also the motion of any object falling freely to the surface of the Earth.

Partly as the result of Newton's work—which could very easily have seemed useless to the people of his time—our technological level has grown to what it is in our day. It was he who first derived the laws of force and motion that are utilized by every engineer in his study and work. And it is the engineer who designs our roads, buildings, radios, and television. He designs better equipment for the doctor to help us with our physical needs and rockets to take us into space. All these things he can do partly because Newton, an astronomer, wanted to understand something purely for the sake of understanding.

Pure science is an essential part of a society as long as that society tries to progress and better itself. A pure scientist is like an explorer who charts unknown territories. He desires only to learn what is unknown; he has no thought of producing further gadgetry but uses the benefits of technology and repays society by giving a new knowledge, a new understanding. To him pure science is enjoyable because it is the quest for understanding.

Astronomy has not always been a pure science. In its beginnings the celestial objects were thought to be gods who intimately affected the people's lives. It was noticed, for example, that the seasons, which meant so much to the agrarian peoples of 3000 B.C., could be correlated with the motions of the sun. Accordingly the sun, and the moon too, played an important role in ancient astronomy.

As the sun stays in the sky for a longer period each day the weather becomes warmer. It is time to plant the crops. To the people who lived in ancient Egypt the lengthening day meant something more: the Nile, the source of their livelihood, would overflow its banks and supply water to the ground so that their crops could grow.

To be able to predict when the Nile would overflow became a vital concern of the ancient Egyptians. Their priests noticed that when a very bright star, the one we call Sirius, rose above the horizon concurrently with the sun, they could expect the river to overflow within a matter of days. The importance of the Nile in their lives led the priests to study more carefully the motions of the sun.

They found that the sun seemed to travel not only across the sky, but through the field of stars as well. If the sun rose with Sirius one morning, it would rise a little later than Sirius the next. After about 91 days Sirius would be high in the sky when the sun appeared on the eastern horizon. After about 182 days Sirius would be setting in the west when the sun was just rising above the horizon in the east. After 365 days, however, the sun would again rise with Sirius, the Nile would overflow, and a new year would begin. Thus the calendar was born.

The month was born of the motions of the moon. There is a new moon every 29½ days, which made it difficult to establish a year with an integral number of full months. This could not and cannot be done simply because 29½ does not divide into 365 an integral number of times.

When Sirius and the sun rise concurrently they appear at different points on the eastern horizon, Sirius more to the southeast than the sun. Huge temples were built with long narrow corridors directed to the exact spot where Sirius would appear. These dark corridors eliminated most of the light of the dawn and enabled the priests to see Sirius more clearly. Through prolonged observation, coupled with this improved method, they came to realize that Sirius rose concurrently with the sun not once every 365 days but every 365¼ days.

If the priests set up a calendar with only 365 days for every year, in 4 years the sun would rise with Sirius one day later than their calendar predicted; in 8 years it would be 2 days late. After 100 years the Nile would overflow it banks 25 days later than the calendar date which had been set 100 years previously.

The priests realized that in order to correct this error in the length of the year they would have to add one day every fourth year and thus make their predictions more accurate. The calendar itself, however, was not changed until the time of Julius Caesar when the Romans took over this knowledge and officially adopted leap year.

In the course of the ensuing centuries it was found that the simple leap year overcorrected the calendar since the year is actually 365.2422 days long or a trifle less than 365¼ days. This inaccuracy of the Julian Calendar, as it is called, was corrected in the sixteenth century by Pope Gregory XIII. The Gregorian Calendar, the one we use today, drops leap year at the close of each century excepting every fourth century. The century year is not a leap year unless it is divisible by 400.

The Egyptians had other troubles, too. After many years of observation they found that Sirius could no longer be used for predicting the overflow of the Nile. The year stayed the same length, but something had happened that made Sirius and the other stars appear to move very slowly in the sky. As a result, the priests had to find other stars that rose concurrently with the sun when the Nile was to overflow and new temples were built to observe these stars.

The Egyptians were well aware of the four cardinal points of the compass (North, South, East, and West). They noticed that each night the stars seemed to travel through the sky in circles centered on a common point in the north (frontispiece). For many years there was a star near that point, but this star did not stay there through the centuries; it moved just as Sirius seemed to move in the sky. After hundreds of years another star was at that special place in the sky, about which the stars seem to rotate.

If we look out into the sky at night we can locate this spot, the north celestial pole, because there is a fairly bright star near it. This star is now called the North Star, or Polaris. Just as 3000 years ago a different star marked the north pole of the sky, 3000 years from now Polaris will not mark the north celestial pole; in fact there will be no star so close to that point. Polaris, however, will again be the North Star 26,000 years from now. Even before the time of Christ the ancient astronomers knew about this apparent motion of the stars, a motion we now call precession, and whose cycle lasts not 24 hours nor 365¼ days, but 26,000 years.

The brighter stars were named by the ancient peoples, chiefly the Arabs and the Greeks. Such names as Aldebaran, Betelgeuse, Deneb, and Vega are Arabic in origin. Others, such as Antares, Canopus, Procyon, and Pollux, were given to us by the Greeks. A few, such as Capella, were derived from the Latin.

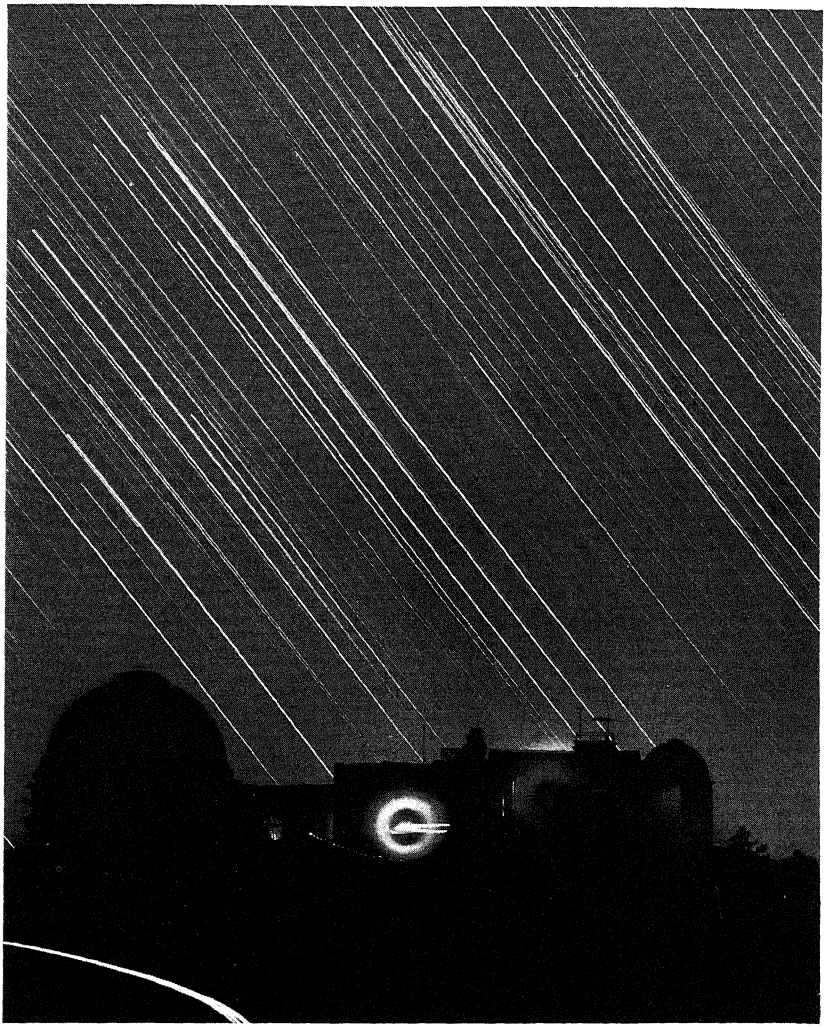


Figure 1-1 Diurnal motion of the stars near the celestial equator appear almost as straight lines. (Lick Observatory)

These ancient peoples watched the sky closely and told stories about the figures they imagined to be represented by the configurations of stars. These figures, our constellations, divide our sky as counties divide a state. Examples are the Big Dipper (correctly called Ursa Major, the Big Bear); Orion, the hunter; Gemini, the twins (from which we derive our expression "by Jiminy"); Scorpius; and the archer, Sagittarius. Most of the constellations seen from the northern hemisphere are the subject of much myth and folklore.

The fainter stars were not named in antiquity but are now named according to the catalog in which they appear. For example, B.D. +30°3639 is a star in the catalog compiled in Bonn called the *Bonner Durchmusterung*. Nonstellar objects are also named after the catalog in which they are listed. M 31 is in Messier's catalog, and N.G.C. 6523 is listed in the New General Catalog. In addition to catalog numbers some of the more famous nonstellar objects have names that derive from the appearance of the object (for example, the Crab nebula, the Owl nebula) or from the constellation in which they appear (for example, the Orion nebula).

We take pride in our particular modern civilization with its many miraculous inventions, but should we not take pride also in the other great civilizations that mankind has produced? Many people in different parts of the world have known a great deal about the very complicated motions of the sun, moon, stars, and planets without the aid of our technology.

In 2254 B.C. the early Chinese civilizations had a calendar that was essentially correct. In A.D. 1279, at the time of Kubla Khan, Chinese astronomers built the first observatory in the world to be equipped with instruments. The Mayas, of what is now Central America, had a calendar that accounted for not only the motions of the sun but also the motions of the moon, Venus, and Mars. This calendar was revised in the year A.D. 1091. The ancient Hindus had developed a calendar that included periods of time up to 4,320,000,000 years. This, interestingly enough, is about the age of the Earth.

Over 2000 years ago the Phoenicians were using the stars to guide their ships on long voyages across the Mediterranean Sea, through the Pillars of Hercules, and up to the British Isles. The Polynesians made even more astounding sea voyages. Unlike the Phoenicians, they had no coast lines which could be used as landmarks. Sometimes the Polynesians sailed between islands more than a thousand miles apart.

This seems amazing to us who tend to look at any but our modern technical society as primitive. We have newspapers, radios, and printed calendars to tell us the day of the month. As a result the average person today is not aware of or concerned with the movements of celestial objects. The sun and stars are not our timepieces;

the clock and the calendar serve us better. But both our calendar and the rate of the clock are determined by the motion of the Earth as it turns on its axis once a day and as it makes its complete trip around the sun each year. Modern man, living in large metropolitan areas ablaze with lights and shrouded with the gases and dust of civilization, scarcely sees the stars on the *celestial sphere* over his head. He hardly glances at the beauty of the celestial sphere, the myriads of stars, and the graceful motions of the planets. He is aware only of the sun and occasionally of the moon, but he is not aware that both the sun and the moon move on the celestial sphere—the dome of the heavens.

Astronomy was a practical science in its early conception, when it was needed to navigate and to develop a calendar. In 1609, however, the approach to astronomy changed. It was then that Galileo first looked through his homemade telescope and saw the mountains on the moon, as well as four other moons revolving about Jupiter. He noticed that Mars and Saturn are not like the stars, for when they are magnified they appear as disks, while the stars remain but points of light. He saw that Venus is different from the other planets in that it goes through phases like the moon. Galileo was fascinated by these phenomena, and his interest in them for their own sake helped turn astronomy into a pure science.

Galileo was concerned with things so distant, so vast, and so mysterious that he followed the urge to learn about our universe even though what he learned might have no direct effect upon men's lives. He and the other astronomers who have followed have seen the unseen; they are revealing mysteries that seem almost beyond the comprehension of mankind.

Yet studying the vastness of space and time may have some effect, not on our daily life perhaps, but upon our thinking and concept of life. Little by little we realize that the Earth, once thought to be the very center of the entire universe, is really just a bit of matter revolving around one of billions of stars that form, together with vast volumes of cosmic dust and gas, a huge galaxy. Comprehending this, we must next confront the fact that this huge galaxy of stars, which we call the Milky Way system, is only one of billions of such galaxies, each with its billions of stars. We are lost in the universe so long as we speak in terms of miles and not light years; so long as we look with our eyes and not with huge telescopes; so long as we think of mankind as having control of vast amounts of energy when he controls the nucleus of the atom, and fail to compare this energy with that released by an exploding star.

Perhaps this concept of the universe will, after all, have some effect upon our daily lives: that Earth, with powerful mankind scurrying over its surface, is less than insignificant in size and influence when measured on a universal scale.

For Further Reading

Bernhard, Herbert J., et al., *New Handbook of the Heavens,* McGraw-Hill paperback, New York, 1964.

Menzel, Donald H., *A Field Guide to the Stars and Planets,* Houghton Mifflin Co., Boston, 1964.

Ronan, Colin, *The Astronomers,* Hill and Wang, New York, 1964.

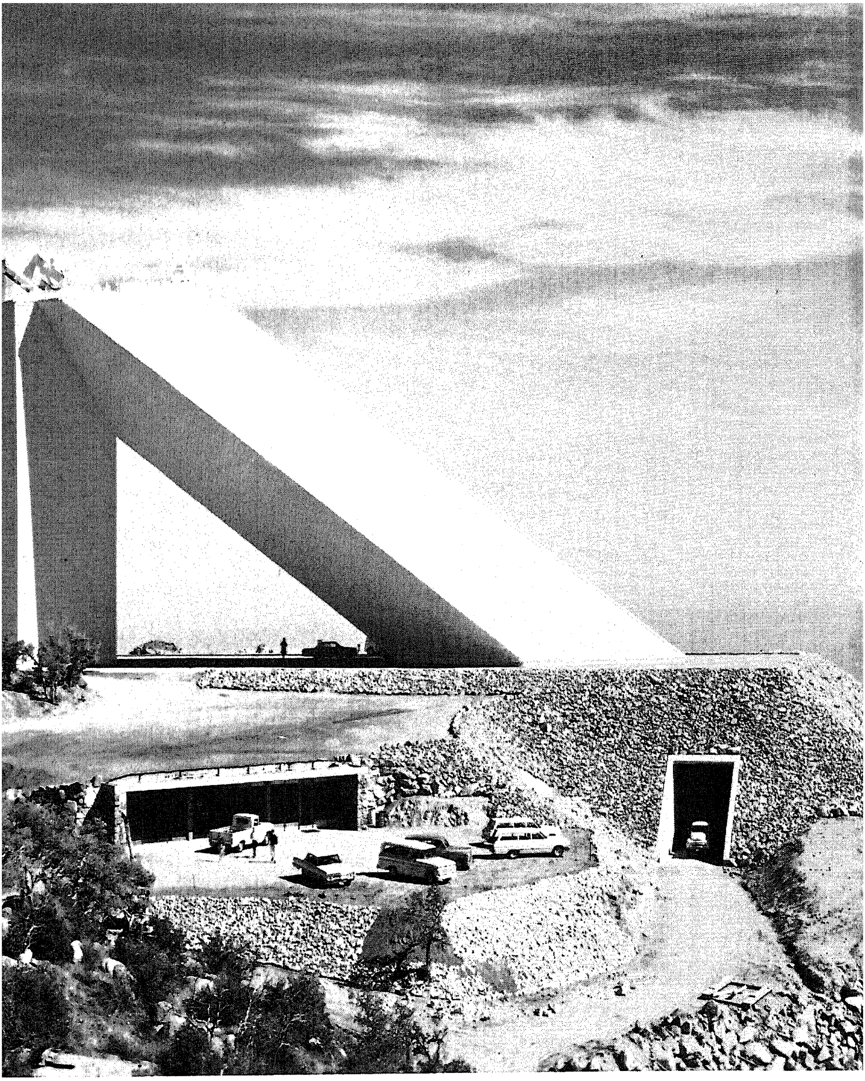
Thiel, Rudolf, *And There Was Light,* Mentor, New York, 1960

Beidleman, R. G., "Light Reflections in the Winter Sky," *Natural History,* p. 38, February 1961.

Price, D. J. de Solla, "Unworldly Mechanics," *Natural History,* p. 9, March 1962.

Price, D. J. de Solla, "Astronomy's Past Preserved at Jaipur," *Natural History,* p. 48, June–July 1964.

"The Sun," *Natural History,* p. 38, April 1963.



McMath Solar Telescope, Kitt Peak National Observatory, Arizona. (Kitt Peak National Observatory)

Chapter 2 *Basic Tools and Methods*

The physicist, the geologist, the biologist, and the chemist are all able to study their subjects in the laboratory or at least to go into the field and make direct observations by many different methods. The astronomer, on the other hand, studies objects that are too big and too far away. With the exception of meteorites (bits of material that move in orbits about the sun much as the Earth does and that occasionally fall onto the Earth), astronomical objects can be studied mainly by the light that we receive from them.

Since the astronomer is dependent upon light we must first consider how it is possible to learn so much about the stars, planets, and other astronomical objects from it alone. To understand this we will need to know some of the characteristics of light. These will be taken up one at a time.

The first characteristic of light is that in free space it travels in a straight line. If we assume this to be true, then a star's position

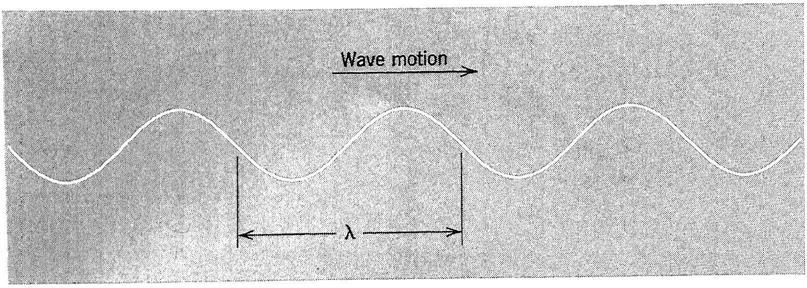


Figure 2-1

in the sky can be determined by noting the direction from which the light comes; that is, by pointing the telescope to the star. Although this may sound very simple, we shall see in later chapters that it is possible for complications to arise even at this point.

Next, many experiments led us to believe that light has wave-like characteristics; in some respects it behaves like a wave caused by a pebble dropped into a smooth pond. A series of these waves on a surface of water is called a wave train (Figure 2-1). Each wave in that train is made up of one crest and one trough. The length of this crest and trough is called the *wavelength.* Waves on the surface of a pond might have a length of 1 centimeter (1 inch equals 2.54 centimeters, or 1 centimeter (cm) equals about 0.4 inch). The wavelength of light is much shorter.

Although it is difficult to measure, the wavelengths of light have been measured with great precision. The wavelength of violet light is about 4.0×10^{-5} cm* (about 1.6×10^{-5} in.); the wavelength of red light is about 7.0×10^{-5} cm (about 2.8×10^{-5} in.). The wavelengths of all the other colors lie between these two extremes. Since the wavelengths of light are so short a unit of measure, the *angstrom* (abbreviated Å) is generally used. One angstrom equals 10^{-8} cm, so the wavelength of violet light is about 4,000 Å; the wavelength of red light is about 7,000 Å.

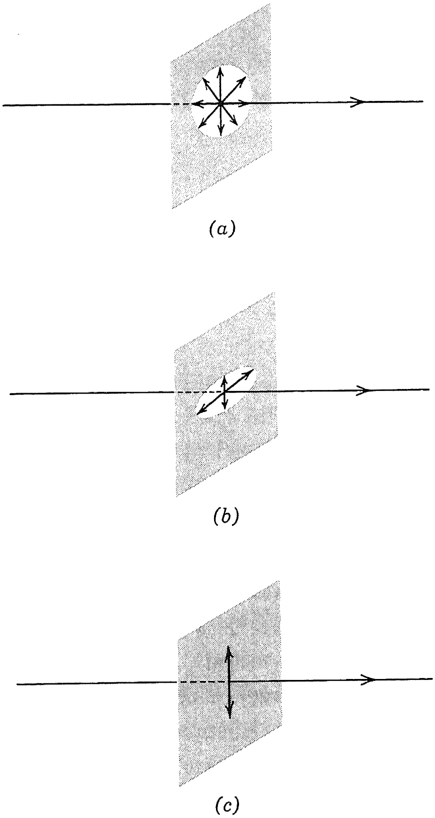
---

* When dealing with very large and very small numbers, it is much easier to use the exponential system for keeping track of the number of zeros. This system is used throughout this book:

$1 \times 10^{-10} = 0.000,000,000,1$
$4.0 \times 10^{-5} = 0.000,04$
$2.1 \times 10^{-2} = 0.021$
$10^{-1} = 0.1$
$10^{0} = 1.0$
$3 \times 10^{1} = 30$
$6.1 \times 10^{2} = 610$
$10^{3} = 1000$
$4.0 \times 10^{6} = 4,000,000$
$5.5 \times 10^{9} = 5,500,000,000$

The surface of the water on which a wave train passes moves at right angles to the direction of wave travel—but only up and down, not right and left. The wave travels along the surface; a cork on the surface will bob up and down as the wave passes.

If the medium through which the wave moves (such as water) vibrates in a direction at right angles to the line of wave travel, the wave is called a *transverse wave*. Light is a transverse wave, but it vibrates in all directions which are perpendicular to the direc-


(a)

(b)

(c)

Figure 2-2 (*a*) Unpolarized light; (*b*) partially polarized light; (*c*) plane polarized light.

tion of wave travel. A line can be perpendicular to another line, or it can be perpendicular to a plane. A line perpendicular to a plane is perpendicular to all lines in that plane passing through the point of intersection of the plane and line of wave travel (Figure 2-2*a*).

If the direction of vibration within this plane is restricted, the wave is said to be partially polarized, that is, to vibrate in one direction more than another (Figure 2-2*b*). A wave on the surface of water is completely polarized, for it vibrates in only one direction (Figure 2-2*c*). Light emitted from an incandescent lamp is unpolarized, and light from the sun is unpolarized. Light reflecting from a surface is partially polarized; the extent of polarization depends upon the angle of reflection and the material from which the light reflects. Consequently, studies of the extent of polarization of light reflected from the surfaces of the moon and of the planets give us some idea of the nature of their surfaces.

Velocity of Light

Light travels at a very high velocity; a velocity so high that it becomes of interest to learn how it was first measured. Before it was measured, no one knew whether the velocity of light was finite or infinite. If it were finite, light would require a certain length of time to travel from one point to another. If the velocity were infinite, however, light would need no time at all to travel from one point to another.

The first measurement was made as early as 1675 by an astronomer named Roemer, who at the time was studying the motions of the satellites of Jupiter. Each satellite goes around Jupiter in a set length of time, called a *period*. The period of each satellite had been determined and Roemer was trying to make a more accurate determination. He timed one of the satellites from the moment it emerged from the shadow of Jupiter until it emerged once again, the satellite having made one complete revolution during this time.

Roemer realized that he would be able to determine a period more accurately if he timed many periods and took their average. This involved him in observations over a considerable length of time, in the course of which a curious phenomenon came to his attention. For the first six months the satellite seemed to emerge from Jupiter's shadow somewhat later than he had anticipated. Then for the second six months it emerged a little earlier until it was back on schedule again.

Roemer explained this by a hypothesis that combined the motion

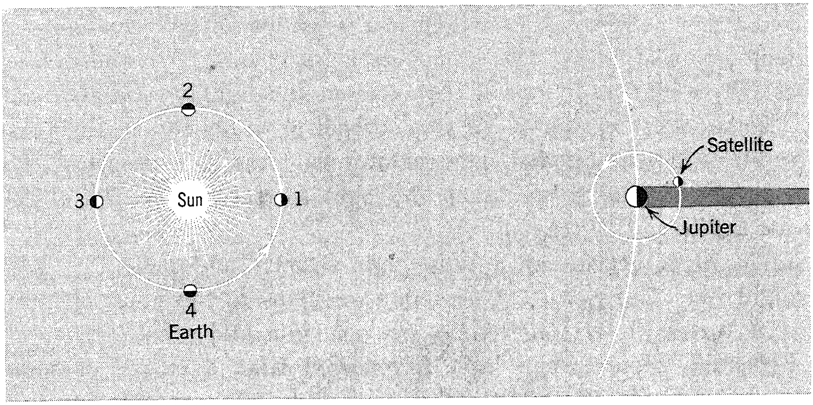


Figure 2-3 The relative positions of the Earth, Jupiter, and one of Jupiter's satellites used by Roemer to measure the velocity of light.

of the Earth around the sun with the assumption that light must have a finite velocity (Figure 2-3). When the Earth was in position 1, he obtained a good value for the length of the period. But as the Earth moved on to position 2 and then to position 3 the satellite appeared to emerge a little later because the light that it reflected and by which alone it could be observed had to travel farther each time to reach the Earth as the Earth receded from Jupiter. (Jupiter moves so much more slowly than the Earth that we may consider it stationary for purposes of this explanation.) Since the satellite appeared to emerge later each time, the length of its period seemed to be increasing.

As the Earth moved from position 3 to position 4 and back to 1 again, the satellite appeared to emerge from the shadow somewhat earlier because the light had a shorter distance to travel each time. Thus during the second six months the period appeared to get shorter until the Earth had returned to the same relative position between the sun and Jupiter, and the period was again what it had been at the beginning of the study.

Roemer knew the approximate diameter of the Earth's orbit and assuming that the period of the satellite did not actually change but only appeared so on the moving Earth, he was able to calculate the velocity of light. The value that he obtained was fairly good; his main source of error was an inaccurate value for the diameter of the Earth's orbit. The velocity now accepted is 186,284 miles per second.

From this discussion it is evident that when we observe any celestial object we observe not the object itself as it is at the moment of observation, but as it was at the moment the light we observe left it. That is, what we see is the object itself at a moment in the past. The interval between that moment and the present is measured by the length of time it has taken the light to travel from the object to the Earth.

Light, then, travels at a very high velocity and generally in a straight line. The two most obvious exceptions to the generalization that it travels in a straight line are *reflection* and *refraction.*

Reflection occurs when light is reflected from a mirror or some other object; refraction, when light travels from air into some other substance, such as glass or water. This can be seen when a pencil is placed in a glass of water. The pencil appears to bend at the surface of the water. That the pencil itself is not bent can be shown by removing it. It is the light bending as it emerges from the water that creates the illusion that the pencil is bent. Refraction results from a change of velocity when light passes from one medium into another. If its velocity is less in the second medium, the direction of travel is bent toward the perpendicular to the surface. Light travels more rapidly in air than in glass, so it bends toward the perpendicular when passing from air into glass. Conversely it bends away from the perpendicular when it passes from glass into air.

Only because we can change the direction of light's travel are we able to make telescopes. These two ways of changing the direction of travel enable us to build two basically different types of telescopes.

The Refracting Telescope

The first telescope was apparently built by a Dutch spectacle-maker, Hans Lippershey. This telescope, employing lenses, was a *refracting telescope*. On October 2, 1608, Lippershey applied to the government officials for a thirty-year title of rights and pension on his invention, and the government officials recognized its military value. But news of the invention reached Galileo in Italy in May of 1609. Galileo did not learn any of the details of the instrument, but after reviewing the laws of refraction he was able to construct his first telescope.

The most common refracting telescopes are built according to the simplified drawing in Figure 2-4. The main lens of a telescope is called the *objective lens.* It is the function of the objective lens to gather the parallel rays that go to make up a parallel beam of light and to focus them at one point.

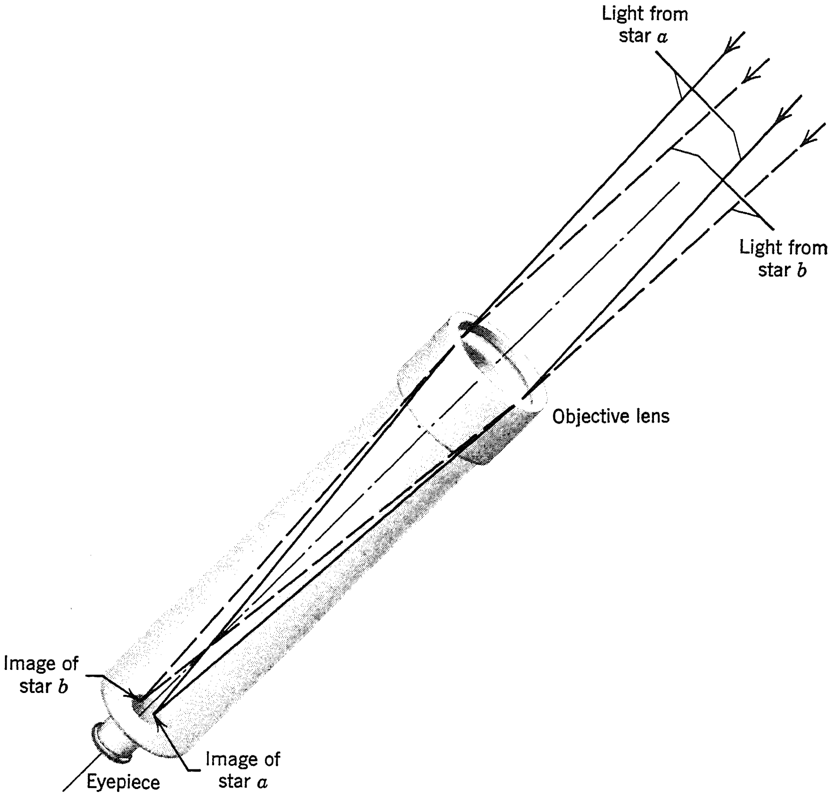


Figure 2-4 Schematic diagram of a refracting telescope.

The light from a single star (*a* or *b*) enters our figure from the right. Since the star is so far away the light travels in an essentially parallel beam. However, the beam of light from star *a* is not parallel to the beam from star *b;* therefore these two beams will be brought to a focus at two different points. All such points from many stars define what is called the *focal plane*. It is possible to place a photographic plate at the focal plane and thus take a picture of a star field. In this case the telescope is essentially nothing more than a large and highly refined camera.

If the telescope is pointed toward a planet, the light from any one point on the planet's surface will be traveling in a parallel beam and all the light from this given point will be brought to sharp focus. This is true for every point on the surface of the planet. All these focused points considered together are called an *image* of the planet. The image is real in the sense that a piece of paper could be placed

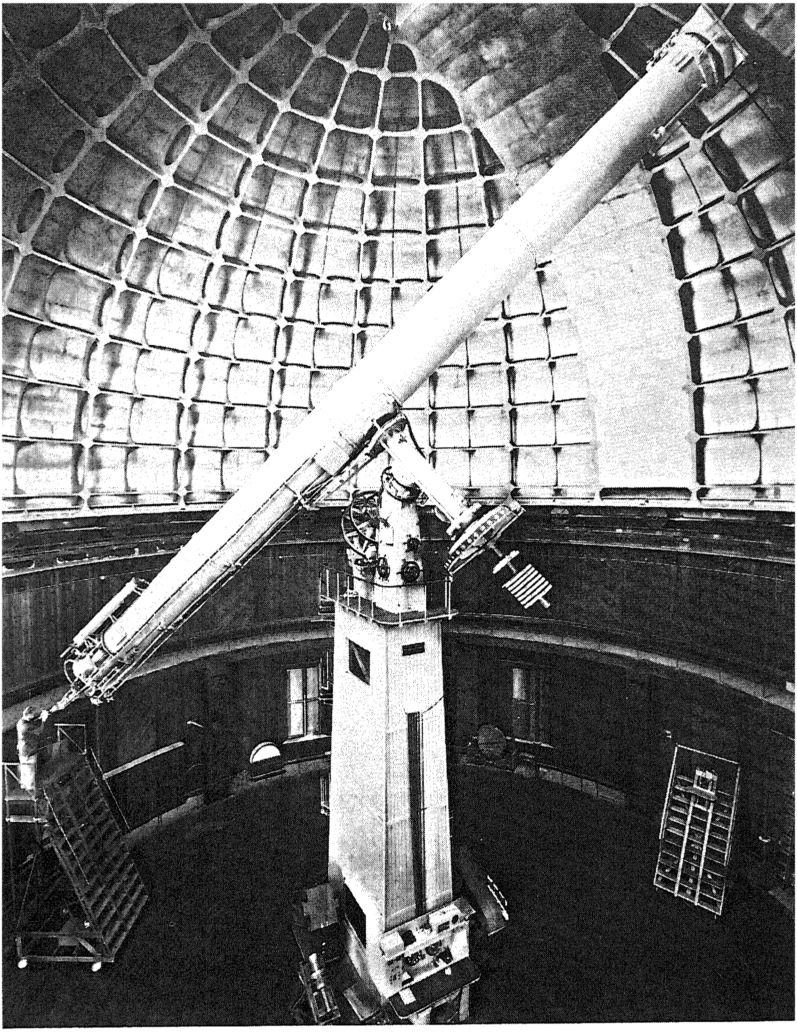


Figure 2-5 The 36-in. refracting telescope of the Lick Observatory, Mount Hamilton. (Lick Observatory)

at the focal plane and the image could be seen on the paper. Again, a photographic plate can be placed at the focal plane and a photograph taken of the planet; or an eyepiece, which is essentially a magnifying glass, can be placed just beyond the focal plane and the image, thus magnified, may be seen by the eye. You cannot use a photographic plate and see the image at the same time.

The larger the objective, the more light it can gather and thus the fainter the object that can be seen or photographed. It is because of this that astronomers have built larger and larger telescopes and have seen correspondingly fainter stars. The three biggest refracting telescopes in the world today are the one at Yerkes Observatory in Wisconsin which has an objective that is 40 in. in diameter; the one at Lick Observatory in California with an objective 36 in. in diameter (Figure 2-5); and a 33-in. refractor at the Meudon Observatory in Paris, France.

The Reflecting Telescope

We have noted that the direction of light can be changed by reflection as well as by refraction. It was Sir Isaac Newton who, in developing the first reflecting telescope, put this principle of optics to use. A schematic drawing of a reflecting telescope is shown in Figure 2-6. Light coming in a parallel beam from a distant star strikes the concave mirror and is converged and focused to a point, *F*. This point is called the *prime focus*. In order to look into an eyepiece

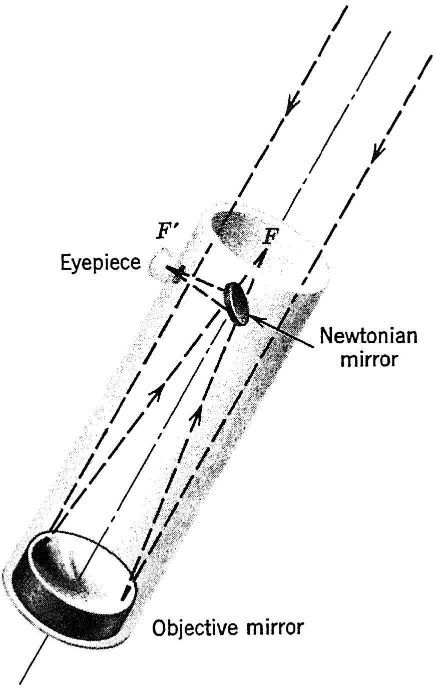


Figure 2-6 Schematic diagram of a reflecting telescope.

placed near the prime focus of this telescope, however, the head of the observer must obstruct the incoming light and thus reduce the amount of light. So Newton placed a small flat mirror in the beam of light converged by the concave mirror and reflected to a point, F', called the *Newtonian focus.* The Newtonian focus is outside the telescope tube and therefore readily available to the astronomer. It is true that even this small mirror prevents some light from reaching the objective, but since it is comparatively small little light is lost.

There are many large reflecting telescopes. The two largest are so large that instead of a small flat mirror they have a little cage where the observer rides at the prime focus. One with a mirror 200 in. in diameter is on Mount Palomar, California; the other is the 120-in. reflector at the Lick Observatory on Mount Hamilton (Figure 2-7). The 102-in. in Crimea, U.S.S.R., is the third largest of the reflecting telescopes.

It is evident that many reflecting telescopes are a good deal larger than the biggest refractors. There are good reasons for this. To bring the light to a sharp focus, the lens or mirror must be ground and polished very accurately. Since the lens can be supported only at the circumference, it bends under its own weight as the telescope is placed in different positions. If the diameter of the lens is greater than about 40 in., this bending becomes excessive, the surface of the lens becomes distorted too much, and the image of the star will not be sharply focused. A mirror, on the other hand, can be supported not only on the edge (as the lens) but also on the back, so bending under its own weight can be prevented. Furthermore, since the light must pass *through* the glass of a lens, that glass must be more nearly perfect throughout. No bubbles or uneven streaks can be tolerated. The mirror, on the other hand, has a thin coat of aluminum on the concave surface to give better reflection and less loss of light. Light does not pass through the glass, in fact never even reaches it, and consequently the glass does not have to be as nearly perfect except over its concave surface. Furthermore, there are at least two surfaces to be ground and polished on a lens and only one surface on a mirror; and since grinding and polishing is a long tedious job, the mirror is easier to build.

If the light does not even touch the glass of the mirror, why must telescopic mirrors be made of glass? There are several reasons. Since glass is amorphous and not crystalline and since it can be made homogeneous throughout, its surface can be ground and polished more

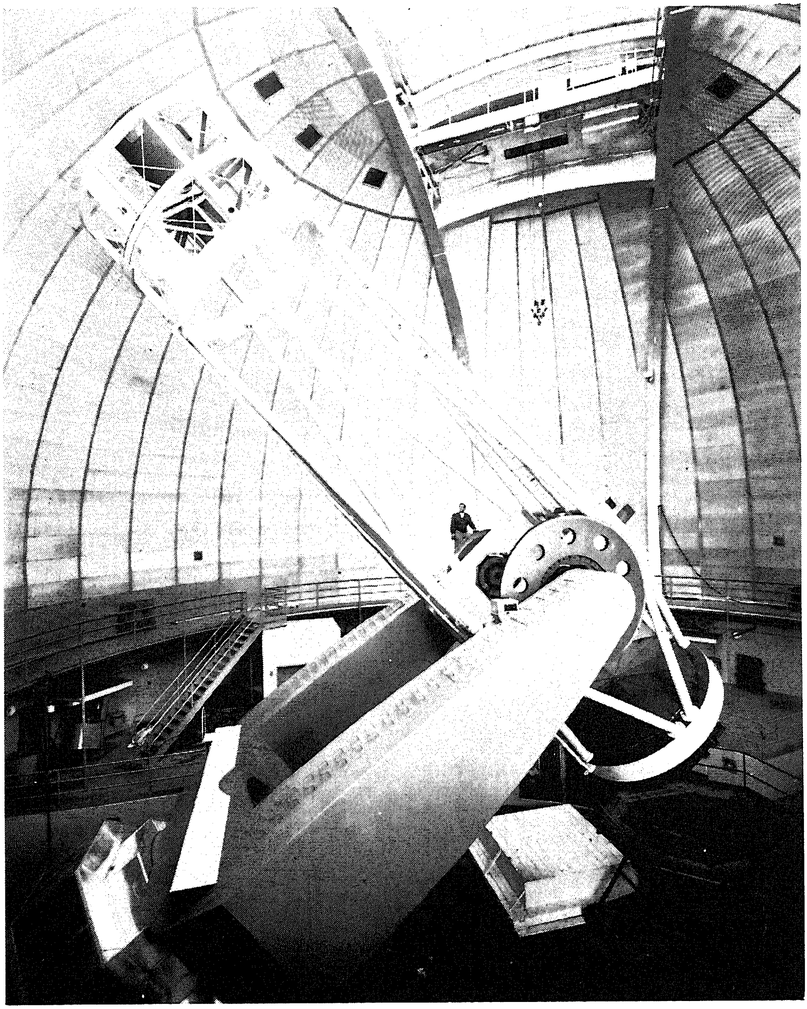


Figure 2-7 The 120-in. reflecting telescope at the Lick Observatory, Mount Hamilton. (Lick Observatory)

accurately than that of most other substances. Another factor is that glass, especially Pyrex, expands and contracts less with changes in temperature than most other substances. This is important not only because expansion and contraction change the mirror's *focal length,* the distance between it and its prime focus; but more important, expansion and contraction change the shape of the mirror's surface causing the image to become blurred.

If the shape of the concave surface upon which the aluminum coating is placed is a portion of a sphere, the stellar images will not be sharp points of light. This failure of a spherical mirror is called *spherical aberration,* and is corrected by making the concave surface parabolic in shape. (For the definition of a parabola see p. 144.)

Chromatic Aberration

There is another reason why the mirror is preferred in most cases. Unfortunately, refraction of light depends on the color of light so that blue light is brought to a focus which is closer to the lens than the focus for red light. This is seen in Figure 2-8*a.* If we place an eyepiece at the focus for the blue light, the star will have a blue center with a red ring around it. If we photograph the star and place the film at the blue focus, the picture we take will not be a sharp point but a blurred image. This phenomenon is called *chromatic aberration* and is a great annoyance to astronomers. Fortunately it can be corrected somewhat by placing a second lens directly behind the primary lens. If this second lens is concave (one that diverges a parallel beam of light) and made of a different kind of glass, it will cause the beam of light to converge less rapidly; it causes the blue light in the beam to converge even less rapidly than the red light so that the foci of all colors are brought more closely together (Figure 2-8*b*). Of course an additional lens increases the cost of the telescope.

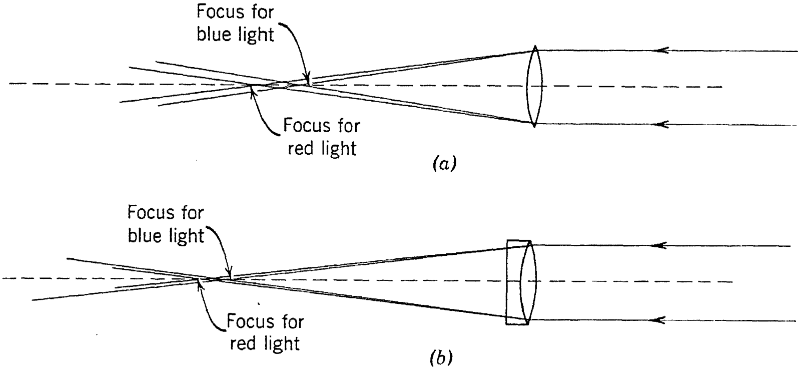


Figure 2-8 (*a*) Chromatic aberration results from the fact that a single lens will not focus the light of all colors to the same point. (*b*) With the addition of a second lens chromatic aberration can be partially corrected.

The correction is never perfect; the refracting telescope always has some chromatic aberration, for only two colors can be brought to the same focus by this two-lens combination. Reflection, however, does not depend on the color of the light; consequently, all colors are brought together at one focus.

The Schmidt Telescope

Very little astronomical research is done visually, since the eye can be deceived and the memory can be faulty. If we photograph the stars we then have a permanent record that can be studied at any later time. Since astronomers work with very faint light it is desirable to build a telescope that will enable us to photograph very faint objects without too long an exposure. A special type of telescope, called the Schmidt in honor of its inventor, can not only cut down the exposure time but also includes a larger portion of the sky in one photograph. A photograph from a Schmidt telescope may cover a portion of the sky about 7° on each side, whereas a photograph from a reflecting telescope may cover only about 1° on each side.

The Schmidt telescope, invented in 1930, uses a spherical mirror to reflect the light to the prime focus. However, since a spherical mirror causes spherical aberration, Schmidt introduced a *correcting plate* at the upper opening of the telescope tube to correct the aberration.

Each of these major types of telescopes, the refractor, the reflector and the Schmidt, has certain advantages over the others. If the surface detail of the planets is to be studied, a large refractor is best. If a larger portion of the sky is to be taken with a relatively short exposure time, a Schmidt telescope is used. If a very faint object is to be examined in detail then a large reflector should be utilized.

To compare the photograph taken by a Schmidt and a large reflector, compare the photographs in Figure 2-9. Photograph *a* was taken with the 18-in. Schmidt on Mount Palomar and photograph *b* was taken with the 100-in. reflector on Mount Wilson. These are photographs of the Orion nebula.

The telescope is the astronomer's basic tool and with it he can take pictures of objects in the sky. But the light that arrives from such distant objects can tell us more than just what they look like and where they are located. Again we must refer to Newton's work, which provides the basis for so much astronomy.

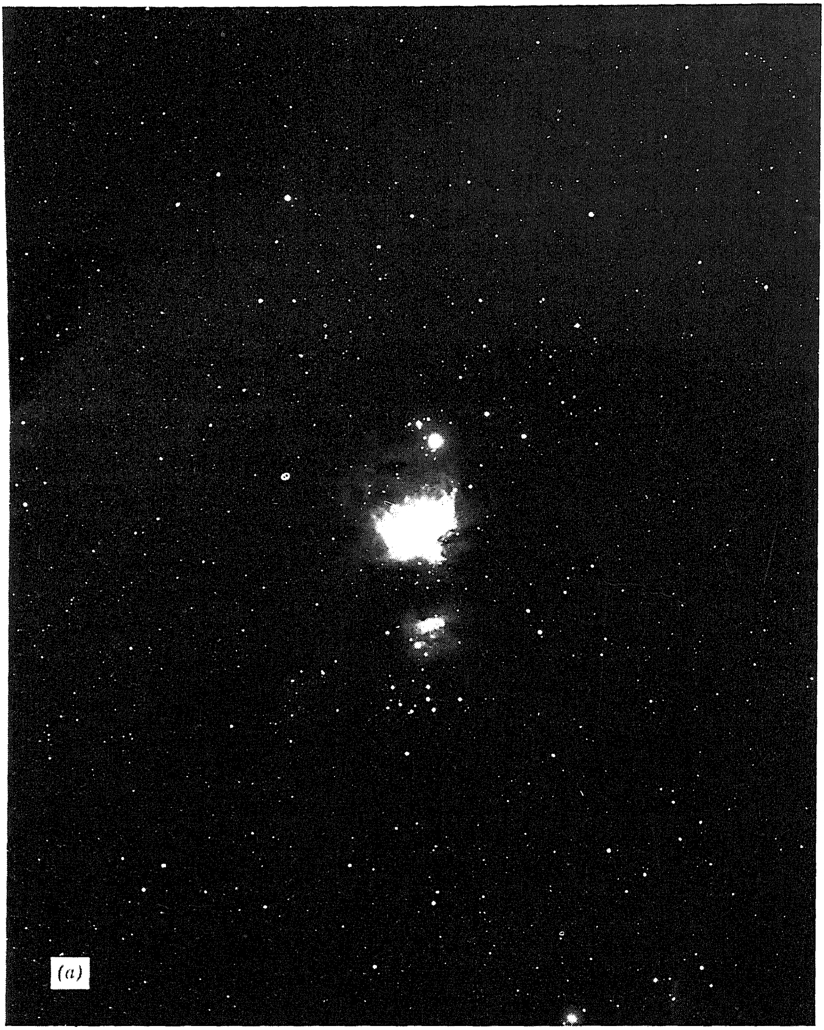


Figure 2-9*a* **The Orion nebula as photographed with 18-in. Schmidt camera on Mount Palomar. (Mount Wilson and Palomar Observatories)**

The Spectrograph

In discussing refraction we found that the angle through which light is refracted depends on the color of the light. Newton found that by passing light through a prism instead of a lens not only would the direction of the light's travel be changed but it would also be broken up into various colors. In the case of the lens we called this

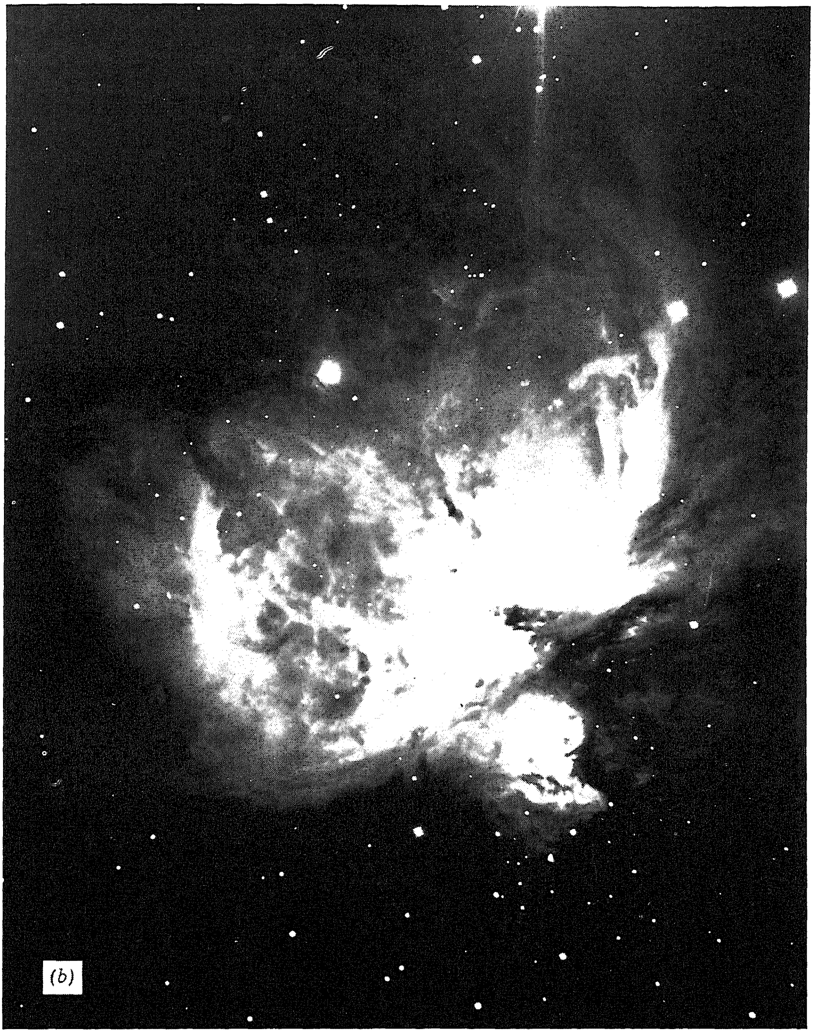


Figure 2-9*b* **The Orion nebula as photographed with the 100-in. telescope on Mount Wilson. (Mount Wilson and Palomar Observatories)**

chromatic aberration and considered it a nuisance. But this nuisance is the basis for one of science's very useful tools.

Light from the sun and all celestial objects is not a single color but a mixture. The light from most stars is nearly white. A prism breaks up white light into its component colors, or *spectrum*. A rainbow is a spectrum, the raindrops act as prisms. In studying the spectrum of the sun, another astronomer, Fraunhofer, found that if he

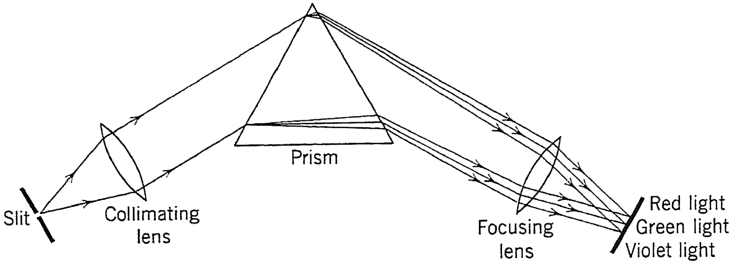


Figure 2-10 A schematic diagram of a spectrograph.

put the image of the sun on a narrow slit in front of the prism the spectrum contained a number of dark lines. A narrow slit can be made by placing the cutting edge of two razor blades very close together. This will allow only a very narrow beam of light to pass through (Figure 2-10). The narrow beam will diverge until it passes through a first lens, called the *collimator.* It then becomes a parallel beam of light. The parallel beam is sent through a prism, which disperses the light into its component colors. It then travels through another lens by which it is focused on a photographic plate. Thus the spectrum is photographed. The entire instrument is called a *spectrograph.*

Each dark line that Fraunhofer found is an image of the narrow slit. If he had placed the original beam of light on a small circular hole he would have found small dark circular spots on the spectrum rather than dark lines. If the slit is made wider, the lines become wider. Fraunhofer did not know the cause of these dark lines but he did letter them *A, B, C,* etc., according to their position in the spectrum. This lettering is still used. It remained for another man, Kirchhoff, to discover what in the sun caused the dark lines.

Spectra

Kirchhoff found that a solid object like red-hot iron will give only a continuous spectrum, that is, a continuous array of colors from violet through blue, green, yellow, and orange to red. An incandescent gas under pressure or an incandescent liquid will also emit a continuous spectrum. On the other hand a gas under low pressure, that is, at much less than atmospheric pressure, when caused to emit light (as in a neon sign) will not give a continuous spectrum but a series of bright lines on a dark background. Such a spectrum is called a *bright line spectrum* and is seen in Figure 2-11*a*. Here again each

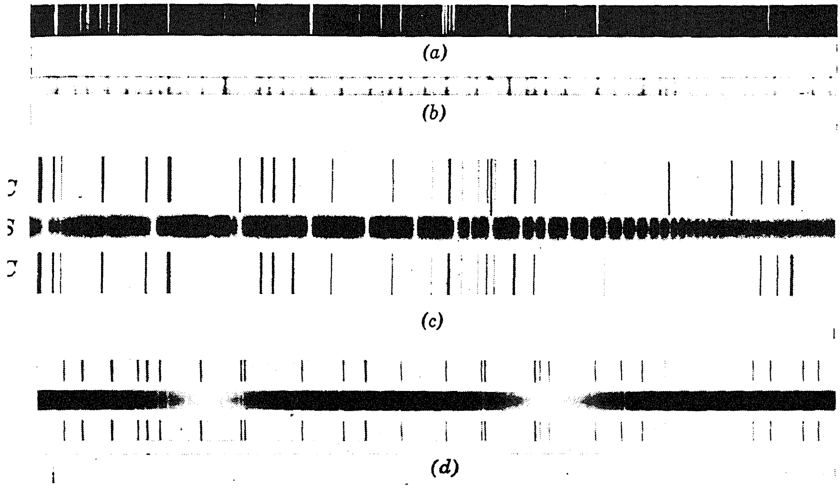


Figure 2-11 (*a*) A bright line spectrum of iron. (Lick Observatory) (*b*) An absorption spectrum of a star showing some iron lines. (Lick Observatory) (*c*) A stellar spectrum *S* showing a portion of the hydrogen spectrum and the comparison spectra *C* above and below the stellar spectrum. Both (*c*) and (*d*) are reproductions of negatives and consequently the bright comparison lines are dark, and the absorption lines are light. (Mount Wilson and Palomar Observatories) (*d*) The stellar iron lines are displaced slightly toward the left (violet) of the iron comparison lines. (Mount Wilson and Palomar Observatories)

line is an image of the slit. If a circular hole had been used, a series of differently colored round spots would have been seen. It is easier, however, to work with thin lines.

Each gas has its own distinctive spectrum composed of a certain number of lines arranged in a fixed sequence. The spectrum of hydrogen (Figure 2-11*c*) is quite simple, whereas that caused by iron when vaporized (Figure 2-11*a*) has many lines and is quite complex. It may seem difficult to vaporize iron, but all that need be done is to cause electricity to spark between two pieces of iron; the spark vaporizes a bit of the metal and thus gives a bright line spectrum.

If the light from a source that gives a continuous spectrum is allowed to pass through a large container full of a gas that is cooler than the source and under low pressure, the continuous spectrum from the source is crossed by a series of dark lines. These dark lines are in the same position as the bright lines for the same gas when incandescent. This type of spectrum is known as a *dark line* or *absorption spectrum* (Figure 2-11*b*).

It should be noted that these lines can originate only with a gas under low pressure. Even the smallest dust particles between the observer and the source of a continuous spectrum will not yield absorption lines.

The work of Newton, Fraunhofer, and Kirchhoff enables us to determine what chemical elements are in the atmosphere of a star. An atmosphere acts like a relatively cool gas that absorbs the light from a hotter source which emits a continuous spectrum. Thus the sun's atmosphere caused the dark lines that Fraunhofer first saw.

To determine what chemicals go to make up a star's atmosphere we can take a bright line source from a spark between two pieces of iron and let its light travel through the spectrograph parallel to the star's light. This gives the comparison spectra shown in Figure 2-11*c* and *d*. The star's spectrum (an absorption spectrum) is in the middle while the two comparison spectra (both bright line spectra) are on the top and bottom. We know that the lines in the comparison spectra are caused by iron and we can see that there are lines in the star's spectrum that are located in nearly the same position. This tells us that the star's atmosphere must contain iron.

Since there are more than 100 elements it would not be practicable to form a comparison spectrum for each one of them. There must be an easier means of identifying the dark lines in a stellar spectrum. The wavelengths of the spectral lines of iron can be measured, and by using them as a reference scale we can measure the wavelengths of all the lines in a star's spectrum. The physicists and chemists, too, are interested in the spectra of elements, and they have measured the wavelength of spectral lines for nearly all of them. The astronomer need only take the wavelengths of the star's spectral lines and match them against the known wavelengths of the spectral lines of the elements.

"Measuring" the wavelength of a stellar spectral line consists in determining its *position* in the spectrum with respect to the positions of the lines in the comparison spectra whose wavelengths are known. In practice, however, this is not quite such an easy matter. There are many, many lines in the spectrum of almost every element, and the gases that make up the atmospheres of the stars are not under the same conditions as gases in a laboratory. Consequently, there are always differences that complicate the situation immensely.

There are other difficulties, too. Notice the placement of the stellar iron lines in Figure 2-11*d* with respect to the iron lines in the comparison spectra. The stellar iron lines are all shifted to the left (the violet).

It might be said that the lines we have attributed to iron are not iron lines but something else very much like iron. But there is no other element that has spectral lines in this same sequence, so there must be some other explanation for the shift.

The Doppler Effect

For the explanation let us look to sound, which also travels by wave motion. Although the nature of sound waves is quite different from that of light waves the effect we wish to discuss is very similar in both.

The wavelength of light determines its color; the wavelength of sound determines its pitch. A very high sound has a short wavelength and a low-pitched sound has a long wavelength.

Let us consider the sound of the horn of a car as it comes toward us, goes past, and travels away from us. The horn emits a sound made up of a series of wave crests and troughs that results in a characteristic pitch if the car is standing still. Since the velocity of sound can be considered constant it is not affected by the speed of the car. Let us further consider the crests only, one at a time. The horn will emit a certain crest which will travel towards our ear with the velocity of sound. If the horn is approaching us, it will move a little closer to our ear before it emits the next crest. Consequently, the next crest will reach us a little more quickly than if it were emitted when the car is standing still. Not only are these two crests closer together but all the crests emitted will be closer together as long as the car is approaching us. Since these crests are closer together when they reach our ear, the pitch of the horn will be higher. Nothing need be said about how far away the car may be, for we are concerned not with its distance but with its velocity toward us.

When the car is directly opposite it will be neither approaching us nor receding from us, and the crests will be spaced the same as if the car were standing still. The pitch of the sound will be the same as if both the car and the listener were at rest.

As the car recedes from us it will move a little farther away after it emits each crest, and the crests will be spaced a little farther apart, that is, the wavelength will be a little longer. In this case fewer crests will reach our ear each second, and the pitch will become lower.

The change in pitch, called the *Doppler effect*, is often effectively mimicked by the entertainer who assumes the role of a radio sportscaster at an automobile race when he imitates the sound of the racing cars as each passes his broadcasting booth.

The Doppler effect applied to light means that the color—wavelength—of the light will change if there is relative motion between the source and the observer. If the measurements of wavelength of the spectral lines of a stellar spectrum are all a bit too short, we assume that the star and the earth are approaching each other. The entire spectrum is shifted to the short wavelength, which for the visible part of the spectrum is toward the violet.

If the measurements of the wavelength of the stellar spectral lines are all a bit too long, then we declare that that star and the earth are receding from each other. The entire spectrum is shifted toward the longer wavelength, which for light is toward the red.

The amount which the wavelength changes $\Delta\lambda$ (Δ, Greek delta used to designate a change in some quantity; and λ, Greek lambda, used to represent wavelength) is proportional to the relative velocity between the source (it may be a star or a planet) and the observer. The change in wavelength $\Delta\lambda$ is also proportional to the unaltered wavelength λ. That is, a spectral line with an unshifted wavelength of 8,000 Å will be shifted twice as far as a spectral line with a wavelength of 4,000 Å.

The change in wavelength also depends upon the velocity of light, c. These rather lengthy statements can be incorporated into a short simple proportion:

$$\frac{\Delta\lambda}{\lambda} = \frac{v}{c}$$

For example, if a spectral line with an unshifted wavelength of 4,860 Å (4.86×10^3 Å) is shifted by +0.15 Å, we can determine the relative velocity of the source by using this equation

$$v = (c)\frac{\Delta\lambda}{\lambda}$$

$$v = (1.86 \times 10^5 \text{ mi/sec}) \frac{0.15 \text{ Å}}{4.86 \times 10^3 \text{ Å}}$$

$$v = 57 \text{ mi/sec}$$

where $c = 1.86 \times 10^5$ mi/sec.

With this well in mind, we can interpret the spectrum in Figure

2-11*d*. All the iron lines have been shifted to the violet, and thus the star must be approaching. When this was discovered it was realized what a powerful tool astronomy had acquired, for by means of the *Doppler shift,* as it is called, an astronomer can determine the velocity with which a star was moving away from or toward us. The velocity can be determined because the faster the star moves the more the spectrum is shifted. The motion of a star either away from or toward the Earth is called the *radial velocity* because it is along a radius (in the line of sight) with the Earth at the center.

The spectrum of a star, then, enables the astronomer to determine not only the chemical composition of the star's atmosphere but also its velocity along the line of sight. Yet a star's spectrum, it must be remembered, is nothing more than starlight after it has been dispersed by a spectrograph.

How much light from the star reaches that spectrograph in the first place? Just by looking up at the stars at night we can see that some stars are brighter than others and that all of them are fainter than the sun. If there were not some way of *measuring* the amount of light received, or at least comparing the amount received from one star with that received from another, the astronomer would have little evidence upon which to base his descriptions of the universe.

Photometry

This question leads us into the field of photometry, the measurement of the amount of light. The first method used by astronomers was simply to estimate with the eye, much as we would do if we were to step outside on a clear night. The first astronomer known to have done this was Hipparchus who lived in the second century B.C. He referred to the brightness as *magnitude* and called the brightest stars those of the first magnitude and the faintest those of the sixth magnitude.

As soon as astronomers began to photograph stars, in the second half of the nineteenth century, they found that the brighter stars left a larger area of exposed emulsion than did the fainter ones (Figure 2-12). By arbitrarily establishing certain stars as standards, we can find the brightness, or *apparent magnitude,* of any star by comparing the size and darkness (the astronomer works with negative plates, not with positive prints) of its image on the plate with the size and darkness of the image left by a standard star. This is more accurate than guessing with one's eye, but it involves certain difficul-

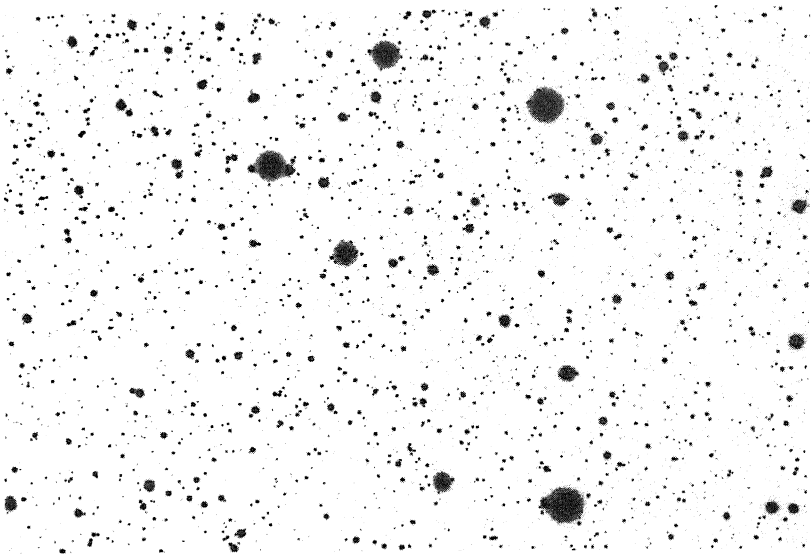


Figure 2-12 The brighter stars leave larger and darker images on the photographic plate than the fainter ones. This photograph, too, is a negative, and thus the stellar images appear dark. (Courtesy of Merle Walker, Lick Observatory)

ties because the photographic emulsion does not respond in the same manner as the eye to the different colors that make up a star's light.

With the advent of electronics, an instrument called a *photocell* has been developed; the photocell permits precise measurements of the brightness of stars. When light strikes a metallic surface, electrons are ejected free of that surface; if these electrons are picked up by a wire which is connected first through a meter—called a *galvanometer*—and then to the metallic surface, the ejected electrons will flow back to the metallic surface through the meter. This flow of electrons constitutes an electric current.

The galvanometer measures the strength of the electric current. The strength of this current is proportional to the number of electrons ejected, which in turn is proportional to the intensity of light. Consequently, the reading of the galvanometer is proportional to the intensity of light.

But even if the photocell is mounted in the focal plane of a large telescope, it receives only a very small amount of light from the very faint stars. Consequently, the electric current is very weak and must be amplified with amplifiers not unlike those used in hi-fi sets.

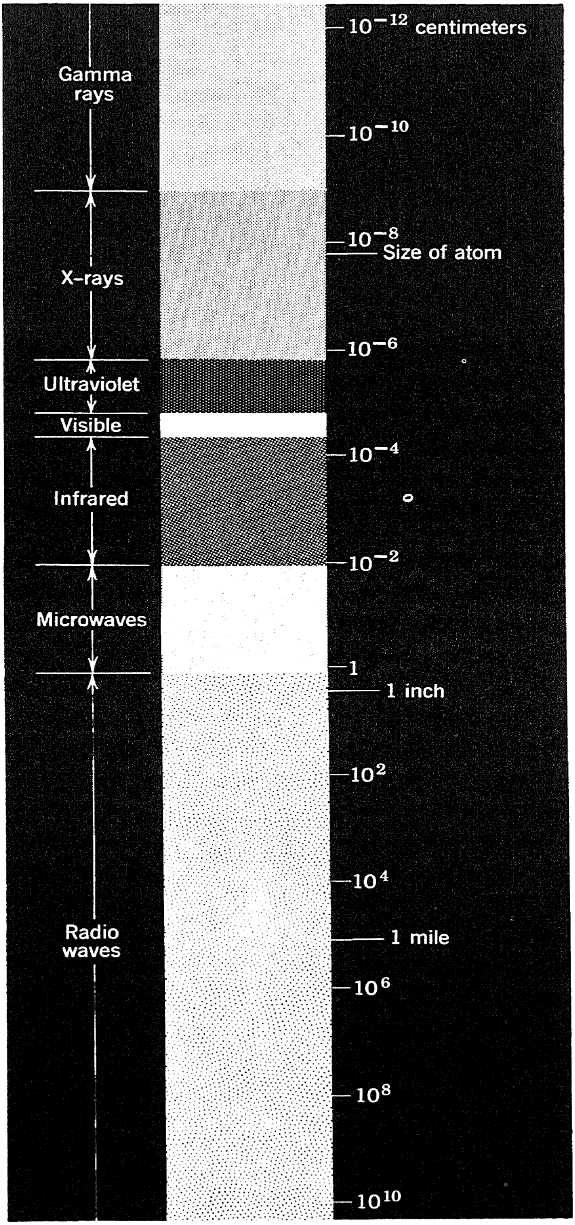


Figure 2-13 The entire electromagnetic spectrum.

Galvanometers can be made very sensitive to small differences in this amplified current, so very precise measurements of differences in the brightness of stars can be made.

If a photocell is used to study the nature of light, the light appears, startlingly enough, to be composed of particles and not waves. These particles are called *photons.* In some experiments light behaves as if it were composed of waves, in other experiments it behaves as if it were composed of particles. Light seems to have a dual nature—a wave-particle nature.

The light that we see is actually only a small part of what is called the *electromagnetic spectrum.* The total electromagnetic spectrum includes radio waves, microwaves, infrared, visible light, ultraviolet, X-rays, and gamma rays. The difference between each of these types of radiation lies in their wavelength; or differences in the energy of the photons. The shorter the wavelength, the greater the energy. Figure 2-13 shows the entire electromagnetic spectrum. The wavelengths are expressed in centimeters.

Substances respond differently to each of these radiations. Glass, for example, is transparent to visible light and yet opaque to other radiations such as the ultraviolet. Quartz, on the other hand, is transparent to more of the ultraviolet. Our atmosphere is transparent to visible light and yet is relatively opaque to the far ultraviolet, that is, the region of ultraviolet that is nearest the X-rays. It is also opaque to X-rays and gamma rays when the radiation comes from outside the atmosphere and must pass through the entire atmosphere. All of these radiations will be transmitted over short distances through the atmosphere. There is a region in the radio waves to which our upper atmosphere (the ionosphere) is quite transparent. This region includes the wavelengths that transmit the FM (frequency-modulated) and television broadcasts. The ionosphere is opaque to the wavelengths of the standard broadcast band, which it reflects back down to the Earth.

Radio Telescopes

In 1931 an employee of the Bell Telephone Laboratories, K. G. Jansky, discovered that celestial objects emit radio waves that can be detected. Not much was done with this startling discovery until the late 1930's when Grote Reber, an electrical engineer from Illinois, built a better instrument to detect radio radiation from the sky. His efforts were stopped by World War II, but out of this war came

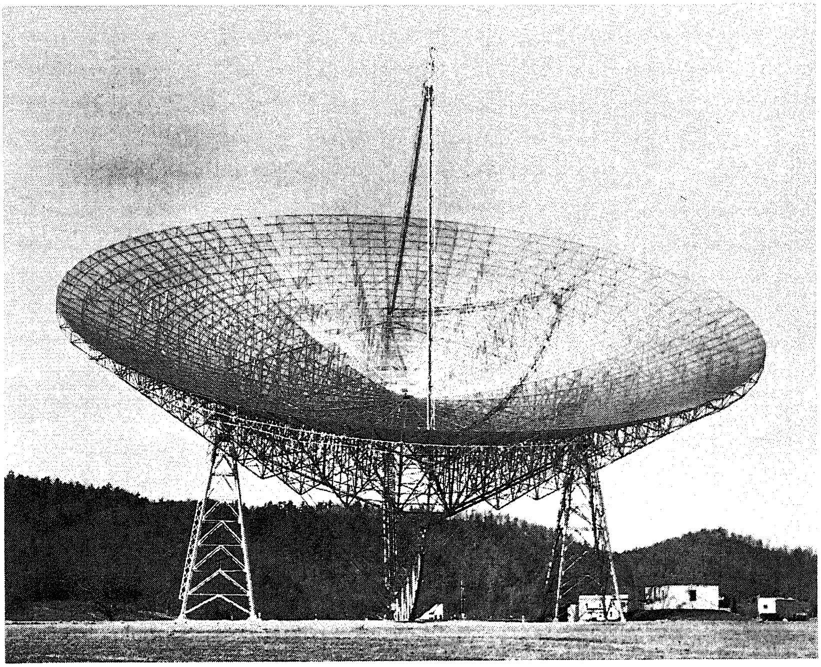


Figure 2-14 The 300-ft radio telescope at Green Bank, West Virginia. (National Radio Astronomy Observatory)

many new techniques of electronics that prepared the way for the present big radio telescopes.

A radio telescope is a reflecting telescope in which the mirror has been replaced by a metallic structure that has a reflecting membrane or mesh made of metal. At the prime focus the eyepiece or photographic plate of the optical telescope has been replaced by an antenna which picks up the radio signals and feeds them into an amplifier, which in turn feeds the signals into a recording device that records the intensity of the radio waves received (Figure 2-14).

A radio telescope has notable advantages and disadvantages. The most important disadvantage is the inability of a radio telescope with a single reflecting surface (as pictured in Figure 2-14) to pinpoint small regions of the sky.

The ability of any telescope to pinpoint a small region of the sky is called *resolving power.* The term is descriptive, because a telescope with high resolving power can separate—or resolve—two objects very close together on the celestial sphere. If stars appear too close together they will be seen not as two stars, but as one. Since distances on the celestial sphere are measured as angles, the separation of objects very close to each other is measured in seconds of arc.* The resolving power of a telescope is defined as the smallest angle separating two stars which can be distinguished as two stars and not merged into one.

Resolving power depends upon the ratio of the wavelength or radiation used to the diameter of the objective of the telescope. Good resolving power is achieved only when the wavelength is small compared to the diameter of the objective. The resolving power expressed as an angle α in seconds of arc is given by the relationship

$$\alpha = 2.1 \times 10^5 \frac{\lambda}{d}$$

where λ is the wavelength of the radiation and d the diameter of the telescope. The effective wavelength of visible light is about 5×10^{-5} cm, so the resolving power of a 4-in. telescope (4 in. = 10 cm) is

$$\alpha = 2.1 \times 10^5 \frac{5 \times 10^{-5} \text{ cm}}{10 \text{ cm}}$$

$$\alpha = 1.0 \text{ seconds of arc}$$

A radio telescope with a diameter of 80 ft (2,400 cm) making an observation with radio waves of wavelength 21 cm has a resolving power of

$$\alpha = 2.1 \times 10^5 \frac{21 \text{ cm}}{2.4 \times 10^3 \text{ cm}}$$

$$\alpha = 1{,}800 \text{ seconds of arc}$$

$$\alpha = 30 \text{ minutes of arc}$$

The moon, as seen from the Earth, has an apparent diameter of 30 minutes of arc therefore, an 80-ft radio telescope observing with 21-cm radio waves could not pinpoint the location of an object on the celestial sphere to an area any smaller than the disk of the moon.

---

* A circle is divided into 360 degrees; the common protractor, a semicircle, has 180 degrees. For small angular measurements, each degree is divided into 60 minutes of arc and each minute of arc is divided into 60 seconds of arc. Consequently there are 3,600 seconds of arc in 1 degree of arc.

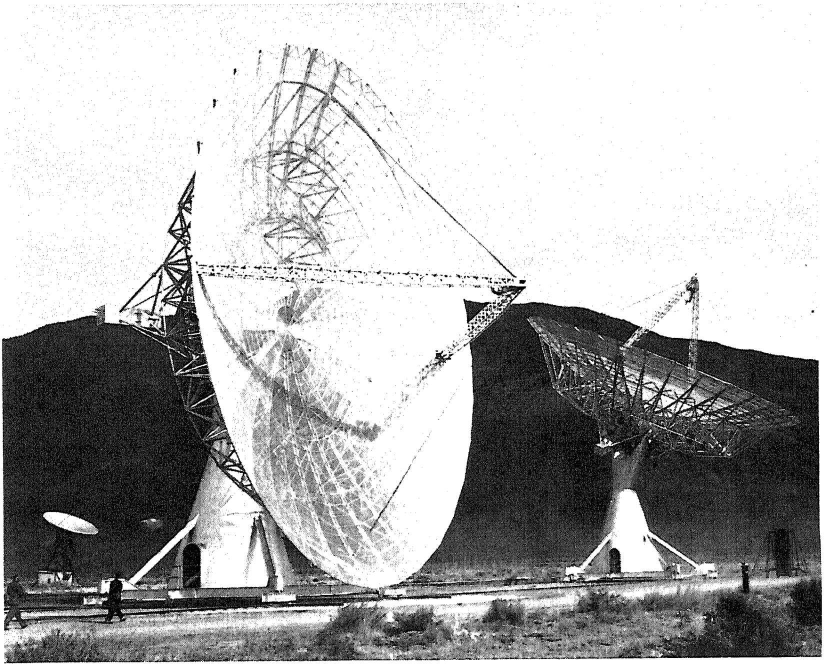


Figure 2-15 The twin radio telescopes used as an interferometer, in Owens Valley, California. (California Institute of Technology)

The resolving power of a radio telescope with one reflecting surface is so poor that ways to overcome this disadvantage were investigated soon after the advantages of the radio telescope became evident. For instance, if a source of radio waves lies near the path of the moon, it is possible to obtain a good idea of its position if the moon passes in front of it and eclipses it. Astronomers have very accurate knowledge of the motion of the moon (its location on the celestial sphere at any given time); therefore, the position of an object emitting radio waves can be determined by observing the time the moon first passes in front of the object and then, as the moon moves, the time it uncovers the object. However, many objects of interest to the radio astronomer will never be eclipsed by the moon.

The limitations of building a radio telescope with a diameter so large that its resolving power would equal that of an optical telescope are obvious; its diameter would have to be 50 miles or more. But it is possible to achieve good resolution by using only portions of an otherwise mammoth radio telescope. Two relatively small telescopes placed ¼ mile apart and electrically connected are as effective in improving the resolving power as one big telescope with a diameter of ¼ mile (Figure 2-15). These two smaller telescopes, however, do not have the ability to gather as much energy as the mammoth telescope.

A system of two telescopes placed some distance apart is called an *interferometer,* because in reality it measures the interference effects when the signals from each of the two radio telescopes are brought together. The resolving power of two radio telescopes placed ¼ mile apart and observing with 21-cm radio waves is close to 2 minutes of arc. The resolving power of the eye to visible light is close to 1 minute of arc! Interferometers have been built in which the antennae are 70 to 80 miles apart.

The radio telescope is a powerful tool with which to investigate the universe and it is a splendid supplement to the other telescopes used by man, who is still largely Earthbound. As long as we continue to use telescopes at the bottom of the atmosphere, however, we must contend with the absorption properties of that atmosphere. Figure 2-16 shows what portions of the electromagnetic spectrum are transmitted by the atmosphere. It can be seen that there is only a relatively small window in the optical region, but there is a much larger window in the radio region. The optical window extends from the near-ultraviolet wavelength about 2,900 Å to about 30,000 Å in the infrared. Thus, the long-wavelength limit of the interval is about 10 times the short-wavelength limit. The radio window, however, extends from a wavelength a little less than 1 cm (¼ in.) to about 3,000 cm (about 100 ft), a factor of more than 3,000. We therefore detect a much larger portion of the electromagnetic spectrum with radio telescopes than we can with optical telescopes. The radio astronomer can scan the radio spectrum by performing an operation entirely equivalent to changing a radio from one station to another or a TV set from one channel to another. The optical astronomer scans the visible part of the spectrum by using a prism.

The most important aspect of radio astronomy, however, is that it gives us a form of observation quite independent from optical observations. The stars, so bright to the eye, are (with the exception of the sun) essentially absent from the sky which the radio telescope "sees." Objects which appear bright to the radio telescope are often times difficult to see with even the biggest optical telescopes. The radio telescopes have increased our power of observation manyfold

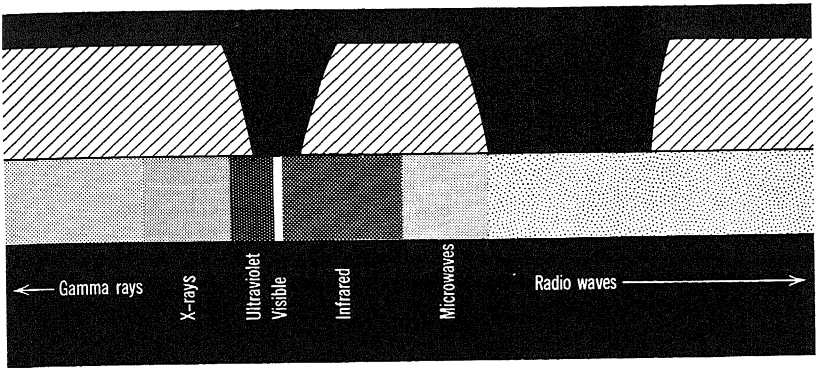


Figure 2-16 Regions of the electromagnetic spectrum absorbed by our atmosphere are shown crosshatched. The resulting "windows" enable us to see out.

and now we have to explain these new observations. Our ideas must catch up with this newly acquired information.

Basic Vocabulary for Subsequent Reading[*]

- Absorption spectrum
- Angstrom
- Apparent magnitude
- Chromatic aberration
- Continuous spectrum
- Doppler effect
- Electromagnetic spectrum
- Emission spectrum
- Focal length
- Interferometer
- Objective
- Prime focus
- Radial velocity
- Resolving power
- Spectral line
- Spectrograph
- Wavelength

For Further Reading

Larmore, Lewis, *Introduction to Photographic Principles,* Dover Publications, New York, 1965.

Mayall, R. N., and Margaret, *Skyshooting; Hunting Stars with Your Camera,* The Ronald Press Co., New York, 1949.

Miczaika, G. R., and W. M. Sinton, *Tools of the Astronomer,* Harvard University Press, Cambridge, Mass., 1961.

[*] A list of words will be given at the end of each chapter to indicate to the reader which words are important for further reading. The student should be very familiar with these words.

Ruechardt, Eduard, *Light, Visible and Invisible,* University of Michigan Press, Ann Arbor, Mich., 1958.

Struve, Otto, and Velta Zebergs, *Astronomy of the 20th Century,* Chapter I, II and VI, Crowell Collier and Macmillan, New York, 1962.

Texereau, Jean, *How to Make a Telescope,* Interscience Publishers, New York, 1957.

Woodbury, David O., *The Glass Giant of Palomar,* Dodd, Mead and Co., New York, 1953.

Brown, D. S., "In-Shop or On-Site Figuring of Large Telescope Mirrors?" *Sky and Telescope,* p. 350, June 1965.

Crawford, David L., "The Kitt Peak 150-inch Telescope," *Sky and Telescope,* p. 268, May 1965.

Findlay, John W., "The 300-foot Radio Telescope at Green Bank," *Sky and Telescope,* p. 68, February 1963.

Friedman, Herbert, "Rocket Astronomy," *Scientific American,* p. 52, June 1959.

Friedman, Herbert, "X-ray Astronomy," *Scientific American,* p. 36, June 1964.

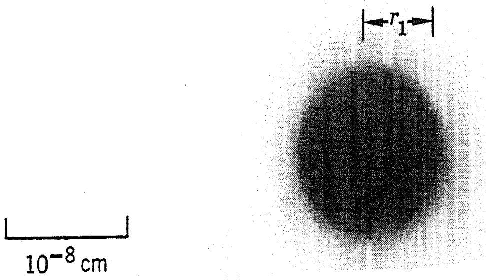
Green, Paul E., and Gordon H. Pettengill, "Exploring the Solar System by Radar," *Sky and Telescope,* p. 9, July 1960.

Hemenway, Paul D., "The Washington 6-inch Transit Circle," *Sky and Telescope,* p. 72, February 1966.

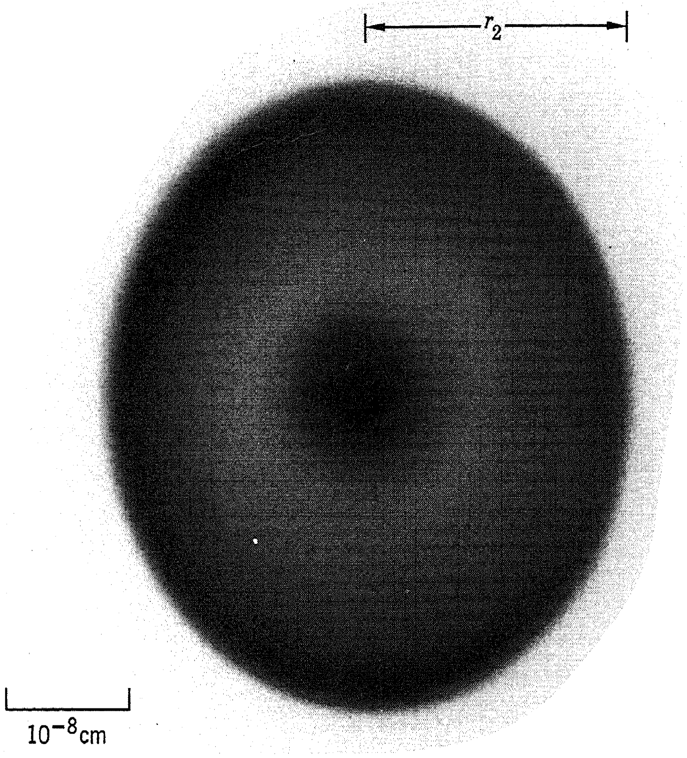
Kiepenheuer, K. O., "The Domeless Solar Refractor of Capri Observatory," *Sky and Telescope,* p. 256, May 1966.

Small, Maxwell, "The New 140-foot Radio Telescope," *Sky and Telescope,* p. 267, November 1965.

"Some Current Programs at Arecibo," *Sky and Telescope,* p. 4, July 1964, and p. 73, August 1964.



(a) $n = 1,\ l = 0,\ m_l = 0$



(b) $n = 2,\ l = 0,\ m_l = 0$

An artist's conception of the hydrogen atom in its (*a*) ground state, and (*b*) with the electron in second energy level. (By permission from *Physics*, K. R. Atkins, John Wiley & Sons, Inc. 1965)

Chapter 3 *The Atom*

The solar system, the stars, and indeed the entire enormous universe are all composed of atoms. Two extremes confront us: that which is too big to comprehend and that which is too small to see.

Atomic Structure

Atoms are composed principally of three fundamental particles: the *proton,* the *neutron,* and the *electron.* Many other particles are found when we investigate the nucleus of the atom, but for the present we shall limit ourselves to a more simplified description.

A mechanical and oversimplified model of the atom is composed of a *nucleus* about which electrons revolve; at the same time the electrons rotate on their axes. This rotation is called *electron spin.* The number of electrons revolving about a particular nucleus (and thus forming an atom of a particular element) depends on that nucleus and determines how the particular atom will react chemically.

The nucleus is by far the most massive part of the atom and is composed, with one exception, of protons and neutrons. That one exception is the simplest atom of all, hydrogen, which has only a proton as its nucleus. Here, as in any atom, the electron is maintained in its orbit about the nucleus by electrical attraction.

There are two kinds of electrical charges, positive and negative. *Like charges* repel each other, whether they are positive or negative. *Unlike charges* attract each other. The proton has a positive charge and the electron a negative one; these unlike charges attract each other and hold the two particles together to form the hydrogen atom. The hydrogen atom is not only the simplest but also the most common atom in the universe.

The next most prevalent atom in the universe, helium, is only slightly more complex; it is the second atom in a sequence based on the number of protons in the nucleus (the *atomic number*). Its nucleus has two protons rather than one. Consequently the helium atom has 2 positive charges and must have two electrons revolving about its nucleus if it is to be electrically neutral.

Since the mass of the proton is 1,836 times that of the electron, the mass of an atom (its *atomic mass*) is given as that of its nucleus. The hydrogen atom has a mass of 1. The most common form of helium, however, has an atomic mass of 4 rather than 2. In addition to the two protons there are two other particles in this atomic nucleus, each of which has a mass about equal to that of the proton. These extra particles are called neutrons because they are electrically neutral, that is, they have neither a positive nor a negative charge. The nucleus of the helium atom, then, contains four particles: two protons each of unit positive charge and mass, and two neutrons of zero charge but of about the same mass as the protons. Two electrons revolve about this nucleus.

There is another variety of helium with an atomic mass of 3. Since it must have two protons (or it would not be helium) it can have but one neutron in its nucleus. These two different types of helium atoms are called *isotopes*, for they have the same (iso) charge on their nucleus and thus fall in the same place (tope) in this sequence, which is based on nuclear charge, even though their masses are different.

The third atom in this sequence is lithium, with three protons in its nucleus and three electrons revolving about it. The most common lithium atom has a mass of 7, which means that there must be four neutrons in its nucleus. There is another isotope of lithium whose mass is 6 and therefore has three neutrons in its nucleus.

The rest of the chemical elements fill out the remainder of the sequence. There are 92 elements found in nature, but by means of particle accelerators (cyclotrons, synchrotrons, etc.) man has created elements that do not occur in nature. Whether they are or were ever formed in nature is another question, but they are not stable and eventually decay radioactively into one of the stable elements. Some natural elements are not stable either, but their rate of decay is so slow that they can still be found on the Earth billions of years after its formation.

This simplified model of the atom is quite easy to imagine; electrons appear as fuzzy bits of negative electricity whirling and spinning about a much more massive nucleus. There are aspects of this oversimplified model, however, which are not supported by observations in physics. Nevertheless, this model is convenient for our purposes and so we will rely on it.

Energy Levels and Light

The idea of the atom is intimately bound to the concept of energy. Consequently, to understand the atom more fully we must first become familiar with the meaning of the term energy. The common usage of the word energy implies the ability to do something. A child has lots of energy; the child is incessantly active. To the scientist, however, a rigorous definition is much more useful. This definition is dependent upon another concept—the concept of work.

The lifting of a weight, the pushing of a car, the accelerating of a rocket, are examples of work—energy is required to do all of these things. Whenever a force is exerted on an object and that object moves as a result of the force, work has been done on that object. We *define work W as the product of the force F exerted and the distance s moved as a result of the force:*

$$W = Fs$$

There is a small catch in this definition, however: both the force and the distance moved must be in the same direction.

Work done on one object increases the energy of that object. The increased energy of that object permits it to do work on a second object. So, if work results in an increase of energy, then energy ought to be able to produce work. In fact, we *define energy as the ability to do work.* Work is action, the act of *changing* the energy of some object; and it is the changes in energy which are significant.

For a 10-lb sack of flour to be lifted from the floor to a greater height, a force must be exerted on it; it then moves the distance to the greater height. Work is done on that sack; its energy is increased. This increase in energy is a consequence of the sack's new position with respect to the Earth; it is energy of position and is called *gravitational potential energy.* The lifting force equals the weight of the sack w, and the distance moved is the height h to which the sack was lifted. Since the work done equals the increase in energy, the change in gravitational potential energy E_p is given mathematically as

$$E_p = wh.$$

The gravitational potential energy would increase by 50 ft-lb if that 10-lb sack were lifted to a height of 5 ft.

An electron and a proton have *electrical potential energy* as a result of their mutual attraction. The amount of potential energy depends upon the relative position of the electron and the proton. In order to separate a proton and an electron, work must be done on them; as a consequence of the work, they will have more energy after being separated than before.

Two electrons, however, repel one another, so work must be done on them to force them closer together. The energy of those electrons will increase as the result of being pushed closer together. Once released, they will fly apart faster after having been pushed closer together. As those electrons fly apart, their electric potential energy decreases. The electrons exert a repulsive force on each other and they move as a result. As the force is exerted they are accelerated. Work is being done on the electrons at the expense of the decrease in potential energy. It would appear, therefore, as if potential energy is being converted into some other form of energy, and since the work done on the electrons increases their velocity, we call this form of energy kinetic energy (Greek, *kinema,* motion, see English *cinema*). *Kinetic energy is defined as energy of motion.* It can be shown that if an object is accelerated from rest, its change in kinetic energy E_k is given by

$$E_k = \tfrac{1}{2}mv^2$$

where *m* is the mass of the object (see Chapter 5) and *v* is its velocity.

This example also serves to demonstrate a very fundamental principle in science: the principle of conservation of energy. There are many forms of energy and one form may be transformed into another. In all such transformations energy is conserved. The total amount of energy before transformation equals the total amount of energy after transformation.

All forms of energy have the ability to do work. Kinetic energy can do work if a moving object hits something. Chemical energy accelerates an automobile as the gasoline explodes in the cylinder head. Electrical energy permits an electric motor to operate a washing machine. Chemical energy produces heat energy and heat energy boils water which makes the steam engine run.

The concept of energy permeates all aspects of science, whether we consider the solar system, the stellar system, or the atom. The laws which describe the motions of the planets and stars, Newton's laws of motion and his law of gravity (see Chapter 5), however, do not apply to the atom. For example, meteorites (small chunks of rock revolving about the sun—see Chapter 7) have both potential and kinetic energy with respect to the Earth. Should a meteorite come close to the Earth, it would fall through the Earth's atmosphere to its surface. Its potential and kinetic energy would be converted to heat. Small meteorites do not survive this generation of heat; they melt and burn in the upper atmosphere.

An isolated and moving electron has both potential and kinetic energy with respect to an isolated proton. Should the electron come under the control of the proton the two would attract each other, but they would not fall together as the meteorite which struck the Earth's surface. Instead, the electron falls to within a certain distance of the proton and then "goes into orbit." The proton and its orbiting electron constitute an atom of hydrogen. The potential and kinetic energy lost in the "fall" of the electron is emitted in the form of light.

The size and shape of this orbit is predetermined by the nucleus. Each atom has many possible orbits, and each orbit represents a definite amount of energy different from that represented by any other orbit. This means that any electron in an atom will have the amount of energy represented by the particular orbit in which it revolves; that is, it occupies a particular level of energy.

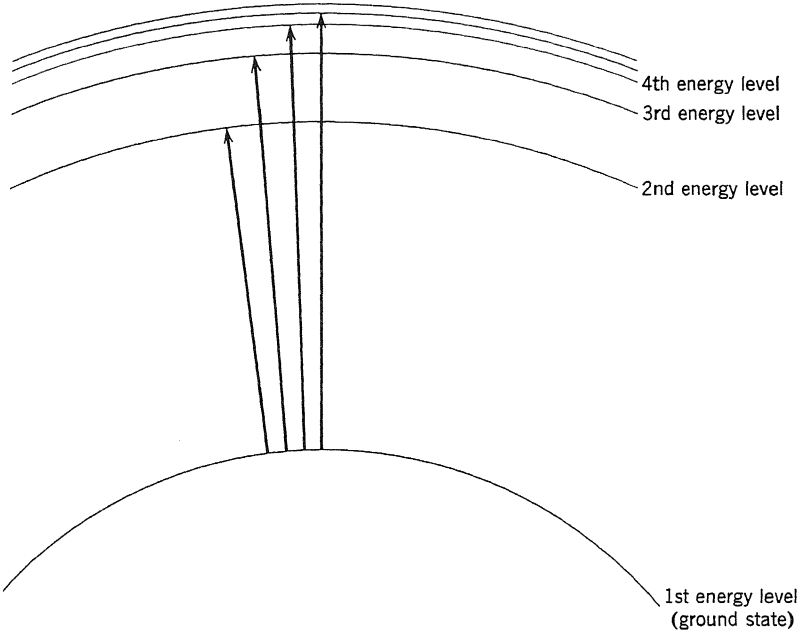
An *energy level* of an atom can accommodate only one electron at a time. Consequently, if to become electrically neutral an atom requires more than one electron (helium needs two; sodium, eleven) then each electron must occupy a different energy level. But each atom has many more energy levels than it has electrons and certain energy levels are more favored by its electrons than others. The energy levels occupied in an undisturbed, electrically neutral atom are thus set in a pattern predetermined by its nucleus.

In general the levels of lowest energy in any particular atom are more favored by the electrons. The levels of lowest energy represent those orbits closest to the nucleus, so the electrons tend to hover as close to the nucleus as they can. For example, the hydrogen atom normally has only one electron. If undisturbed, this electron will seek the lowest energy level, also called the *ground state.*

An electron in an atom may, from an outside source, gain energy that will force it up into a higher energy level. This can be demonstrated by introducing hydrogen into a closed tube at low pressure, and passing an electric current through it. Electricity is a flow of electrons, and moving electrons have kinetic energy. If a free electron with enough energy collides with a hydrogen atom, it will give some of its energy to the orbital electron. The bombarding electron will leave the scene of the collision with a reduced velocity because of its loss of energy. The orbital electron will change to a higher energy level as a result of its increase in energy. A change in energy levels by an orbital electron is called a *transition;* in this example it is an upward transition (Figure 3-1). An atom that has had an electron knocked into one of the higher energy levels is said to be *excited.*

Since the hydrogen atom has many energy levels, to which level does the orbital electron go when bombarded by a free electron? This depends on two factors. The first is the amount of energy of the bombarding electron. The more energy it has the higher the orbital electron can go. But when the electron makes a transition it must go from one energy level to another; it cannot stop between them. Since the energy levels are well defined and furthermore, since the electron is permitted to make a transition only from one energy level to another, the orbital electron can absorb only certain discrete amounts of energy from the bombarding electron, namely, the amount necessary to make a single transition between any two energy levels.

The second factor determining which transition an electron will undergo is the probability of making any given transition. Certain transitions are more likely than others, and this helps determine which discrete amount of energy the orbital electron will absorb. It need



igure 3-1 Electron transitions from the ground state to higher energy levels the hydrogen atom.

ot absorb all the energy made available by the bombarding electron.

The situation is similar to putting a plate on one of a number f shelves in an empty cupboard. The plate must rest on one of e shelves; it cannot be placed between them. To lift the plate from e floor (the ground state) to one of the lower shelves does not equire as much energy as to lift it to one of the higher shelves. f the cupboard has five shelves, one of only five discrete amounts f energy is needed to lift the plate to one of the shelves. To move he plate from the floor to any one shelf, or from one shelf to a igher one, amounts to a transition that the plate cannot make unless t is given precisely the required amount of energy.

It is possible, however, to give the plate a toss into the air so hat it will go higher than the highest shelf. Since there is no shelf bove the cupboard (and we assume no ceiling) it can be tossed vith any amount of energy. Similarly, an electron at any energy level nay be bombarded by a free electron with enough energy to send t above even the highest energy level. When this happens the electron as been knocked free of the nucleus and the atom becomes *ionized.* For this the electron may be given any amount of energy greater than just the amount required for ionization, and since it has gone beyond discrete transitions, it is no longer limited to the absorption of discrete amounts of energy.

To return to the analogy of the plate in the cupboard: if the plate is permitted to fall, it will drop and perhaps break. In striking the floor it gives off all the energy lent it when it was lifted from the floor to one of the higher shelves.

But here the analogy of the plate and shelves becomes incorrect. The plate will remain on the shelf unless disturbed, whereas the electron will fall from the higher energy level of its own accord, usually within 10^{-8} second.

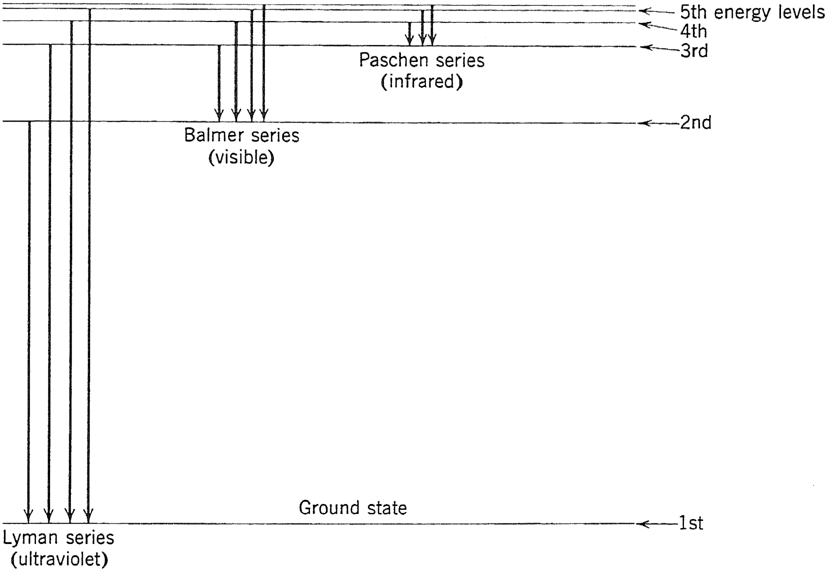
When the electron does fall to one of the lower levels it gives off energy. This energy is radiant energy (visible or ultraviolet light, infrared, radio, etc.) and is emitted as *photons.* A given photon has an amount of energy that depends upon the particular transition through which the electron has fallen. Since an electron in a given atom is permitted only certain transitions in falling, it can give off photons of only certain energies, each representing a particular wavelength of radiant energy. The same transition made downward by electrons in many atoms of the same chemical element gives rise to a bright spectral line.

Many transitions made downward between *different* pairs of energy levels in atoms of the same element produce that element's characteristic bright-line spectrum. We see that in a spectrum each spectral line results from a downward transition between two particular energy levels of an atom.

If the downward transition is short, the amount of energy given off by the electron is small. The photon of radiant energy will travel from the atom with this small amount of energy but regardless of energy, will travel with the velocity of light. The greater the downward transition, the more energy the photon will carry away. Thus the placement of the spectral lines in the spectrum depends on the particular transition made.

The entire electromagnetic spectrum represents a continuous array of wavelengths beginning with the longest, radio waves, and proceeding through the microwaves, infrared, visible light, ultraviolet, and X-ray to gamma rays, the shortest. These wavelengths can be translated into energies, and thus the spectrum is a continuous array of radiant energy. The shorter wavelengths have more energy than the longer; blue light, for instance, has more energy than red.

Whenever an electron makes a transition to the ground state of



igure 3-2 Electron transitions downward to the ground state emit light that ields the spectral sequence called the Lyman series. All downward transitions o the second energy level give rise to the Balmer series; downward transitions to he third energy level yield the Paschen series. The energy levels have been rawn as straight lines for convenience.

ydrogen, for example, there is given off a photon of sufficient energy o place it in the ultraviolet region of the spectrum. All such transiions from any higher energy level to the ground state of hydrogen onstitute a series of spectral lines in the ultraviolet, called the *Lyman eries* (Figure 3-2). If a transition is made down to the second level, he atom emits light in the visible region of the spectrum, and all uch transitions in hydrogen result in a spectral line sequence called he *Balmer series.* Since hydrogen is the most prevalent atom in the niverse and since the Balmer series is in the visible and near-ultraiolet region of the spectrum, this series becomes very important o the astronomer.

If transitions are made from the third to the second energy level f hydrogen, the resulting spectral line is in the red region of the pectrum and has a wavelength of 6,563 Å. Since this is the first line f the Balmer series it is called *Hα* (*H alpha*). If a transition is ade from the fourth to the second energy level, more energy will e given off, and this spectral line, called *Hβ* (*H beta*), is in the lue region with a wavelength of 4,861 Å. The third line in this series results from a transition from the fifth energy level to the second, has a wavelength of 4,340 Å, and is called $H\gamma$ (*H gamma*). There is, theoretically, an infinite number of lines in each of these series; but since each energy level is progressively closer to the preceding one as the outermost level is reached, the lines blend together toward the end of the series (Figure 2-11*c*).

A bright-line or emission spectrum can therefore be explained by electron transitions from one energy level to any other below it. But how about the dark-line spectrum? It was pointed out earlier (see page 29) that a dark-line or absorption spectrum is caused by sending light that consists of all wavelengths (a continuous spectrum) through a low-pressure gas that is cooler than the source of the light. The light in this continuous spectrum consists of photons of all energies within the range of wavelengths emitted. These photons are able to bombard the orbital electrons in the atoms of the gas and cause them to make upward transitions. A given bombarding photon, however, is *completely* absorbed in the process. It disappears as a photon and appears as increased energy in an electron. The electron normally remains in a higher energy level for about a 100-millionth of a second before falling back down to land eventually in the ground state. In making downward transitions, it gives off photons of light which may or may not be of the same wavelength as the energy absorbed in causing the original upward transition. If the electron was originally in the ground state and falls directly back to that state, the photon emitted will be of the same wavelength as that which excited the electron. However, the electron may choose to fall one energy level at a time (*cascading*) and thus the photons emitted will each have less energy (longer wavelength) than the exciting photon. An electron in a higher energy level can be excited to an even higher energy level, but it, too, will eventually wind up in the ground state. If the electron is knocked free of the nucleus by a *photon*, the atom becomes *ionized* and the process is known as *photoionization.*

The absorption spectrum results from this process of absorption and re-emission. But when it happens that the emitted photons are of the same wavelength as the absorbed photons, why should we observe an absorption line?

Let us consider only the original light from the continuous source that is heading straight for our spectrograph (Figure 3-3). All the light would reach our spectrograph were it not for the intervening gas. This gas is composed of atoms that absorb light of only certain wavelengths; *which* wavelengths depends on the transitions possible

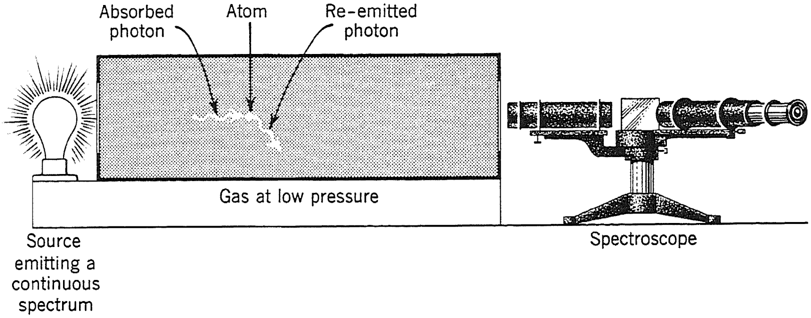


Figure 3-3 White light that shines through a gas cooler than the source of light will have certain colors subtracted from it giving rise to a dark-line spectrum. A photon originally heading for the spectrograph is absorbed by an atom in the gas and re-emitted. The re-emitted photon, however, will probably travel in some other direction and thus will not reach the spectrograph. The result is a dark line in an otherwise continuous spectrum.

in the atoms. The light that is re-emitted as a result of the electrons' downward transitions is re-emitted in *all* directions. Thus only a small portion of the re-emitted light continues in the original direction toward our spectrograph, and consequently an absorption line is dark. If no energy were to reach the spectrograph, such a line would be completely black. The absorption spectrum, therefore, consists of a continuous background from which selected regions (or lines) have been absorbed by the upward transitions of electrons in the atoms that compose the particular intervening gas. The transitions for a given atom, whether they are upward (an absorption line) or downward (an emission line), involve the same amount of energy. Thus the two types of lines resulting from transitions between two given energy levels have the same wavelength; for $H\alpha$ of the Balmer series we can have either an emission or absorption line and each has a wavelength of 6,563 Å.

Thus far we have discussed two methods of exciting an orbital electron. It may be bombarded by another electron, in which case the *only* visible effect produced is the bright lines emitted when the electron falls back down to a lower energy level. Or the electron may be excited by a photon, in which case the visible effect is an absorption line. This is the only manner in which an absorption line can be produced.

The electron may be excited in a third manner—by thermal agitation of the particles composing a hot substance. When iron is heated in a forge or crucible the thermal agitation causes the electrons to make upward transitions. The excited electrons fall back, emitting light which is seen as a red glow. Since this is a solid, or perhaps a liquid, why does it give off a continuous spectrum and not the bright-line spectrum of iron?[*]

To explain this let us imagine some hydrogen inside a closed container that has a piston in one end. When the piston is fully extended the gas is under very low pressure and when excited will emit a bright-line spectrum. If the piston is pushed in, the pressure increases and the energy levels of the hydrogen atoms become distorted by the increased jostling received by the atoms now that they have been forced closer together.

Any electron that makes a transition from a distorted energy level will not give off light of the same wavelength as it would if the energy level were not distorted. Since the energy levels of various atoms are not uniformly distorted, the resulting bright spectral lines become broadened and fuzzy. This is known as *pressure broadening* and provides us with a somewhat simplified explanation of the continuous spectrum. If the pressure of the gas is increased even further the spectral lines eventually become so broadened as to blend into a continuous spectrum. As was noted in Chapter 2, any incandescent solid or liquid emits a continuous spectrum because its atoms (or molecules, for that matter) are close together. Furthermore, each substance emitting a continuous spectrum is opaque.

Basic Vocabulary for Subsequent Reading

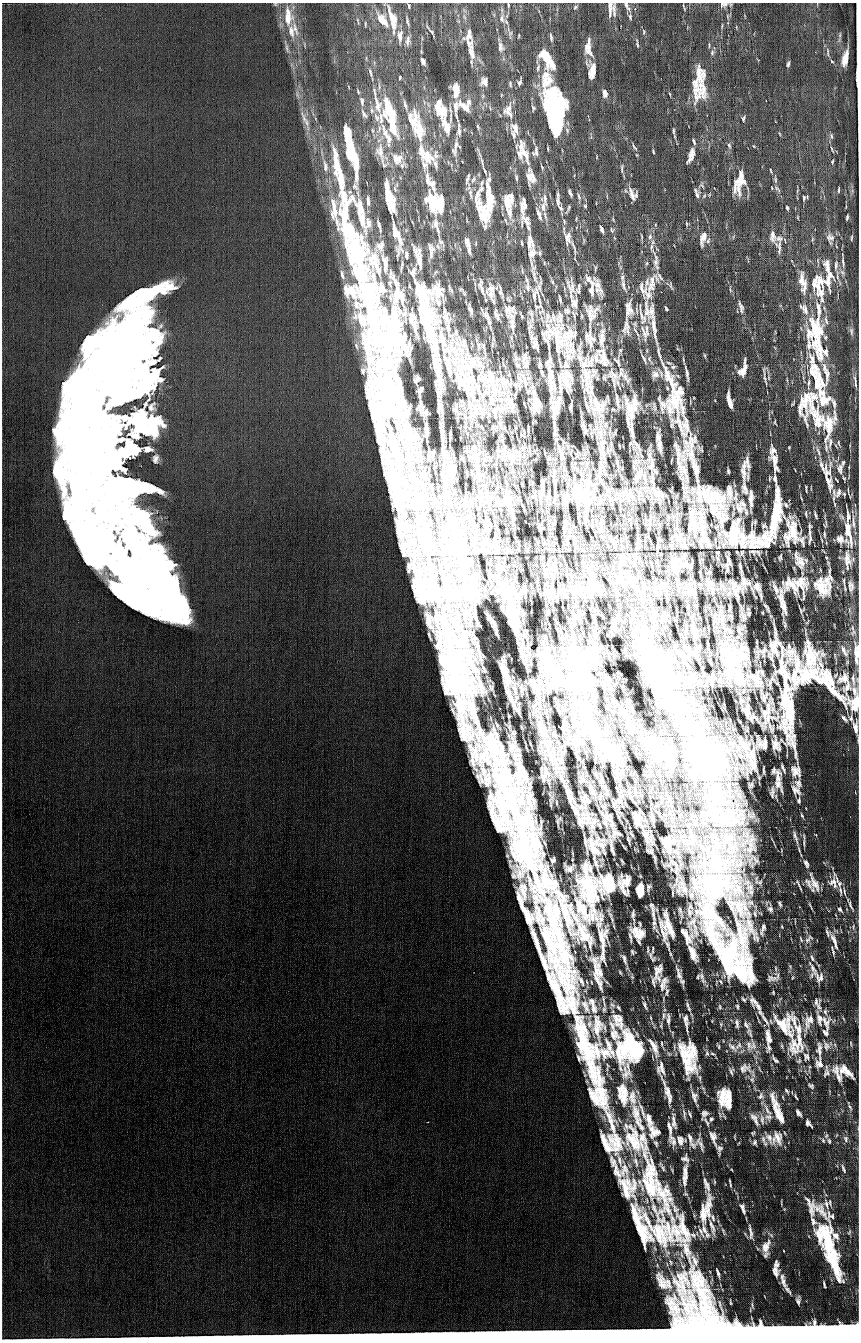
Atomic mass
Atomic number
Balmer series
Electron
Electron spin
Electron transition
Energy level
Ionization
Lyman series
Neutron
Nucleus
Photon
Proton
Pressure broadening

For Further Reading

Ruechardt, Eduard, *Light, Visible and Invisible,* University of Michigan Press, Ann Arbor, Mich., 1958.

Semat, Henry, and Harvey E. White, *Atomic Age Physics,* Holt, Rinehart and Winston, New York, 1959.

[*] See page 28f.



Photograph of the planet Earth taken by Lunar Orbiter I with a portion of the backside of the moon in the right foreground. The east coast of the United States is under the clouds in the upper left portion of the crescent Earth. (NASA)

Chapter 4 *The Physical Characteristics of the Planets*

Now that we have some idea of the astronomer's methods, we can better understand the objects of his studies. To the naked eye all the stars (except the sun) and planets appear as points of light. But in the early telescopes those objects closest to the Earth took on a new appearance. The same is true today: planets appear as disks, whereas the stars appear only as points of light even when viewed with the highest magnification available. The early observations led to the first great change in man's concept of the universe—the idea of a *solar system.*

The solar system is composed of the sun, its largest and central member; the nine planets that revolve about the sun; satellites, comets, asteroids, meteors, and something we might call dust.

The sun is really a star and we shall consider it along with the other stars. For the present we are concerned with the planets, which in the solar system are second only to the sun in importance and size.

Within the solar system, the nine planets have at least two different motions. Each of them *revolves* about the sun in a given length of time just as the satellites of Jupiter revolve about Jupiter. The planet Earth, for instance, takes one year to revolve about the sun; in fact, it is because of this revolution that we have the year.

At the same time each planet *rotates* on its own axis like a merry-go-round. The Earth makes one complete rotation on its axis each day. During part of this rotation you, the observer, are facing the sun; during part of it you are facing away from the sun, in the shadow of the Earth.

All planets describe these two basic motions of revolution and rotation, although with differences. But what of their other physical characteristics? Do planets other than the Earth have atmospheres, soil, conditions that would favor life? Conditions on them are largely determined by their distance from the sun because each, like the Earth, receives its heat from the sun. The closer the planet, the more heat it receives.

Mercury

Mercury, less than half as far from the sun as the Earth, is the planet closest to the sun. To put it more briefly, it is 0.4 A. U. from the sun; A. U. is the abbreviation for a measure of length called the *astronomical unit* which is used by astronomers because the units of length employed on the Earth are all too short. 1 A. U. is equal to the average radius of the Earth's orbit, about 93,000,000 miles.

With a diameter of 3,100 miles and a mass only 0.05 that of the Earth, Mercury is also the smallest planet. It revolves about the sun in 88 days and was long thought to keep the same face toward the sun. However, observations in 1965 made by G. H. Pettengill and R. B. Dyce with the 1,000-ft radio telescope in Puerto Rico by radar (that is, by sending a radio signal to Mercury and observing the echo) indicate that, with respect to the stars, the planet rotates on its axis with a period of about 58.4 days. Consequently, the entire planet receives sunlight; no face is perpetually dark. But since the direction of rotation is the same as the direction of revolution, the period of rotation with respect to the sun is about 175 days. The sun will shine on one point of the surface for a period of time equal to about 87½ days here on Earth.

But how can a radar signal be used to determine the period of rotation of a planet? If the planet rotates relative to the Earth, then ½ of the disk will be rotating toward the Earth and the portion of the signal reflecting from this half of the disk will be received with a Doppler shift toward the shorter wavelengths (higher frequencies). That portion of the signal reflecting from the center of the disk will exhibit no Doppler shift. The portion of the signal reflecting from the half of the disk receding from the Earth will exhibit a Doppler shift to the longer wavelengths (lower frequencies). Unfortunately, very little if any signal reflects from the limb of the planet—the *limb* is the apparent circular edge of the disk. It is the portion of the limb on the equator of the planet which rotates the fastest and thus would cause the greatest Doppler shift.

If the transmitted signal is of only one frequency (in reality a very narrow band of frequencies), then the signal returning from a rotating planet will be spread out over a slightly broader band of frequencies. By comparing the breadth of the band of frequencies of the returning signal with the breadth of the band of the transmitted signal, the speed, and thus the period of rotation, of the planet can be determined.

But the majority of the returning signal reflects from the central part of the disk, and this portion of the signal exhibits little or no Doppler shift. As a result, this portion of the signal masks or drowns out the rest of the signal. Radio astronomers have consequently devised a way to observe only that portion of the signal returning from the part of the disk between the central part and the limb.

That portion of the signal which reflects from the central part of the disk arrives back at Earth first—after all, the planet is nearly spherical and so the central part of the disk is closest to the Earth (Figure 4-1). The signal returning from the regions closer to the limb will be received a moment later. Therefore, if the transmitted signal is a pulse of very short duration, the returning signal will be a pulse of a little longer duration. The first portion of the pulse to arrive must reflect from the center of the disk; the last portions must reflect from near the limb.

By selecting a portion of the returning pulse which arrives a little bit later than the first, the radio astronomer is able to select the signal which returns from a predetermined ring or annulus about the disk of the planet. The Doppler shift caused by the rotation of the planet from this ring about the disk is easier to measure than

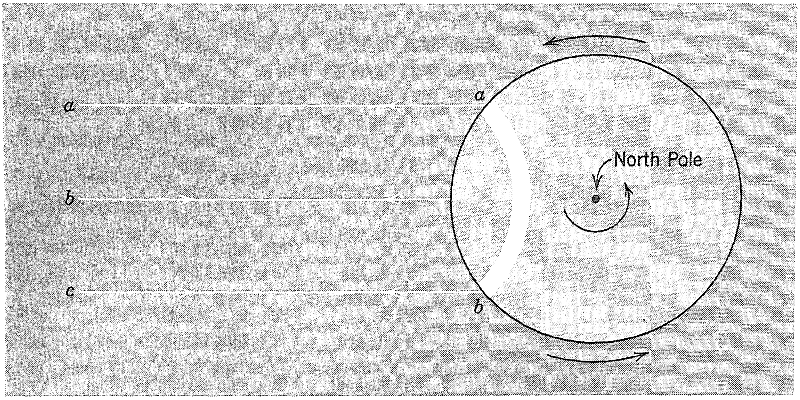


Figure 4-1 If a planet rotates with respect to the Earth, radar waves reflected from the side approaching the Earth (*a*) will be received by our radio telescopes with a Doppler shift to the shorter wavelengths. Radar waves reflecting from the side of the planet receding (*b*) will be received with a shift to the longer wavelengths.

the signal which reflects from the entire disk, because that portion which reflects from the central part has been eliminated.

If the period of rotation of Mercury seems to be settled, the existence of an atmosphere is not entirely certain. Determining if Mercury has an atmosphere may depend in part upon the way the word "atmosphere" is defined. If the term is applied to a layer of gas about a planet which is not only permanent in nature but which is dense enough to be easily detectable, then the answer is an unqualified no. If, however, the term atmosphere is taken to include very rarified gases which may or may not be permanently associated with the planet, then Mercury may have an atmosphere.

Mercury, like every other planet, shines by reflected sunlight. Therefore, its spectrum must be the same as the sun's spectrum—unless it has an atmosphere of its own. (Surface materials do absorb some light but not as spectral lines.) If Mercury has an atmosphere, the sunlight would have to pass through it to reflect from the surface, and then pass back through the atmosphere again before reaching our telescopes. Since this atmosphere would act as a cool gas at low pressure, it would add absorption lines to the solar spectrum. The wavelengths of these lines would indicate the chemical nature of the atmosphere.

There is slight evidence that the spectrum of Mercury does contain spectral lines other than those of the sun. N. A. Kosyrev of the Pulkovo Observatory in Russia reports that he has observed spectroscopic evidence that Mercury does contain a very slight atmosphere of hydrogen gas. But this atmosphere would not remain were it not replenished by protons and electrons (the two constituent parts of the hydrogen atom) from the sun.

Whether a planet can retain an atmosphere or not depends upon the mass and radius of the planet, its temperature, and the kind of gas. The mass and radius of the planet determine the escape velocity at the surface. The *escape velocity* is the velocity an object must achieve in order to escape from that planet. The escape velocity at the surface of Mercury is 2.6 miles per second. (By comparison the escape velocity at the surface of the Earth is 7.0 miles per second.)

The velocity with which the atoms of gas travel depends upon the temperature of the gas and upon the kind of gas. Even if an atmosphere is composed of only one kind of gas, its atoms, because of the atomic collisions, will be traveling at different velocities. As a result of a collision, one atom may acquire a greater velocity, another a lower velocity. The number of atoms at each particular velocity depends upon the temperature of the gas. As a result, the fastest atoms will escape from the planet if they acquire the escape velocity. If the atmosphere contains several kinds of gases, the atoms of the lightest gases will have the greatest velocity. The atoms of those gases with the greatest mass will have the least velocity. Consequently, a planet is more apt to retain gases with more massive atoms; for example, the heaviest gases of all which might comprise a rarified atmosphere are the inert gases krypton and xenon. A planet is more apt to retain oxygen, nitrogen, and carbon dioxide than it is hydrogen and helium, for the latter two gases are the lightest gases of all. It is possible, therefore, that Mercury might have a slight—but a very rarefied—atmosphere composed of the heaviest inert gases.

However, it is well known that the sun expels protons and electrons which, when combined, form atoms of hydrogen. Consequently, it is possible that Mercury retains a slight atmosphere of hydrogen which is always changing; gases leave because the atoms acquire the escape velocity, but these gases are replenished by those expelled by the sun and captured by Mercury. It is also possible that some gases are still erupting from volcanic-like activity on the surface of Mercury.

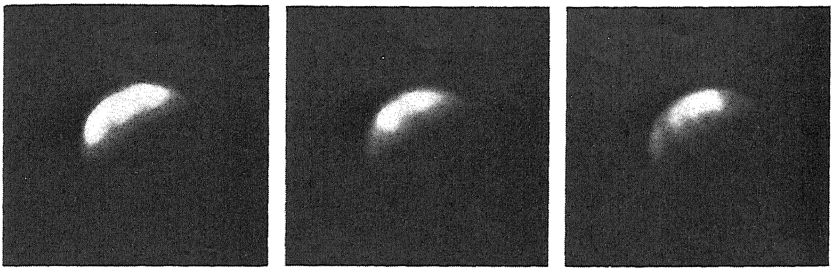
We are led to believe that the surface of Mercury is similar to the surface of the moon (see Chapter 6). Mercury reflects 6% of the light incident on it. This can be determined because we know how much light the sun emits; we know how far Mercury is from the sun and thus how much of the sun's light it intercepts. We also know, at the moment of observation, how far we are from Mercury. With the added knowledge that the intensity of light radiating from the sun and reflecting from Mercury decreases as the square of the distance, we can calculate how much light we on Earth would receive from Mercury if it were a perfect reflector. We receive about 6% of that quantity. Astronomers use the term *albedo* to describe the reflectivity of a surface; the albedo is the fractional part of the incident light which reflects from a surface. The albedo of Mercury is 0.06; the albedo of the moon is 0.07.

The fact that the albedoes of these two bodies are nearly the same and that other properties of the reflected light are also similar (for example, the color and the extent of polarization), leads us to conclude that the surface of Mercury is similar to that of the moon.

The temperature of the surface of Mercury has been measured by a number of methods. Measurements of the temperature of the sunlit side using the infrared region of the electromagnetic spectrum indicate the sunlit face, at its hottest, may reach a temperature of 650°F. More recently, observations of Mercury with radio telescopes using radio waves with a length of about 3.5 cm indicate that the *average* temperature of the sunlit side is not far from 220°F. As expected, the temperature of the surface must drop in those regions where the sunlight strikes the surface more obliquely. Measurements of the nighttime side using radio waves with a length of about 3.4 millimeters (mm) indicates a temperature of about —100°F.

Venus

The second planet from the sun (0.7 A.U.) is Venus. With a diameter of 7,600 miles and a mass 0.89 times that of the Earth, the planet is only a little smaller than the Earth. The optical telescope reveals only a layer of clouds which are opaque to visible light. Occasionally these clouds show markings, although these markings are not permanent (Figure 4-2). Since the markings are only temporary, astronomers cannot determine the period of rotation by watching the planet rotate. The period of rotation has, however, been determined by radar to be about 248 days, but in the direction opposite to the direction of revolution. The planet rotates in *retrograde,* that is, reverse, motion



Aug. 13 Aug. 14 Aug. 15

Figure 4-2*a* Three photographs of Venus taken on consecutive days to show the changes in its cloud cover. (Lick Observatory)

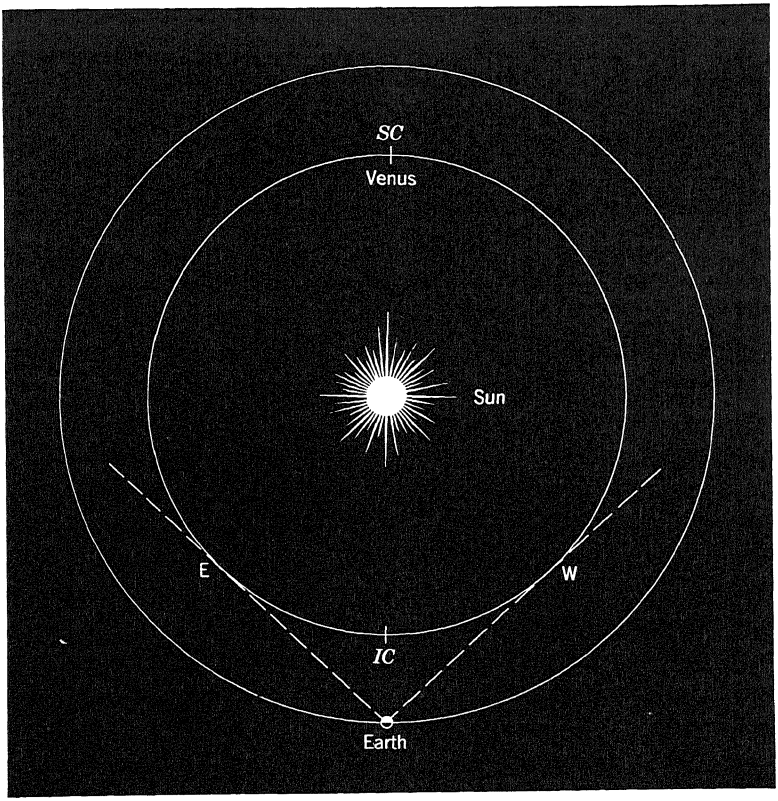


Figure 4-2*b* From the Earth we cannot observe either Mercury or Venus in full phase when either planet is clearly visible and easily observed. At full phase the planets are at *superior conjunction* (*SC*) behind the sun. We can only see a thin crescent at most when either planet is at *inferior conjunction* (*IC*). As viewed from the Earth the planets are farthest from the sun at *eastern elongation* (E), or *western elongation* (W) and are most easily observed at these times. Radar studies can be made near inferior conjunction, however.

The period of rotation with respect to the sun must, therefore, be about 122 days. Consequently, the sun shines on one location for about 61 days.

Retrograde rotation is a bit disconcerting, for it is not the easiest situation to explain. Present theories of the origin and evolution of the solar system (see Chapter 8) attempt to explain only those planets which rotate in the same direction in which they revolve about the sun. At this stage, we simply do not know enough of the past history of the solar system to explain retrograde rotation. We must wait until we have gathered much more information, particularly about the very early history of our solar system. Space exploration appears to be the only method by which we will be able to gather such information.

Since Venus is continually shrouded in clouds we do not have a direct view of the surface. Consequently, all our information about the surface must come from observations made with regions of the electromagnetic spectrum which travel right through the cloud layer. Since radios on Earth can be heard even when it is raining or foggy, radio waves must pass right through clouds. Consequently observations of the surface of Venus are made by observing the radio waves that the planet emits and reflects.

Radio observations indicate that the surface of Venus is very hot. Observations using wavelengths between 2 and 3 cm lead radio astronomers to conclude that the temperature is in excess of 600°F. If this high temperature is of the surface it can be explained by the same atmosphere which prevents us from seeing the surface.

The atmosphere of Venus may well act as a heat trap. The term describing this action is the *greenhouse effect:* it is recognized that radiation of one wavelength—the short wavelengths of visible and infrared light—can travel through the glass of a greenhouse and warm up the interior, but that since the radiations emitted by the interior are of longer wavelength, they are blocked by the glass. An automobile parked in the summer sun with all the windows rolled up demonstrates this principle very well. It is known that the atmosphere of Venus does contain a relatively high percentage of carbon dioxide when compared to the atmosphere of the Earth, and carbon dioxide is one of the more effective producers of the greenhouse effect. The heat trapped by the atmosphere apparently keeps the surface hot.

Venus' atmosphere is also known to contain at least some water vapor. It is difficult to determine how much, because the light that leaves Venus must pass through our own atmosphere, which certainly contains water vapor. It is not easy to differentiate spectral lines caused by the water vapor in our atmosphere from those originating in the atmosphere of Venus.

In 1964, a spectrometer designed and operated by astronomers at Johns Hopkins University was carried aloft by an unmanned balloon to a height of 87,500 ft—above all but 0.03% of the water vapor in the Earth's atmosphere. A photoelectric cell scanned the spectrum formed by the spectrometer in the infrared region, for it is in this region where spectral lines of water vapor appear. These observations indicate that the atmosphere of Venus does indeed contain some water, but that much of it is frozen into ice crystals. The temperature at the top of the layer clouds, as determined by studies in the infrared, is —38°F, so it is not surprising to find the water vapor frozen into ice.

Spectroscopic studies to detect molecular oxygen in the atmosphere of Venus have yielded negative results. There appears to be very, very little oxygen, at least at the top of the cloud layer. It is to be presumed that, like the Earth's atmosphere, the bulk of the atmosphere of Venus is composed of nitrogen gas.

Observations at radio wavelengths in the order of 2 to 5 mm generally yield a temperature lower than 600°F. It may well be that radiations of this wavelength are emitted by a layer of gas in the atmosphere which is colder than the surface. Certainly such layers exist in the Earth's atmosphere.

On December 14, 1962, Mariner II, a spacecraft sent to make observations of Venus, passed within a distance of only 22,000 miles of the planet. A radio telescope aboard the spacecraft made three scans of the planet, recording information and transmitting it to Earth. The information so gained yielded a temperature of 600°F or more, in agreement with Earthbound radio telescopes. But more important, those observations indicate that at this temperature the limb of the planet is cooler than the central regions. This relative coolness of the limb leads to the conclusion that it is the surface which is 600°F and not a layer in the atmosphere.

Observations of the limb are made through a longer path of the atmosphere than observations of the center of the disk. Consequently, observations of the limb are more influenced by the atmosphere above the surface which may well be cooler than the surface.

The radar observations of Venus indicate that the surface of the planet is perhaps as rough as the moon's surface; much remains to be learned about Venus, however, and a great deal of effort is being made in that direction. Radio observations, radar observations, and future space probes will, it is hoped, help penetrate the curtain of clouds surrounding this planet.

Earth

The third planet out from the sun is the Earth. It is special to us because it is our means of life, and it is to our present knowledge unique because it does have life. As our investigations of the universe become more precise and more penetrating, however, it is logical to expect that we will find many, many stars with planets about them and many of these may well have life on them. We have advanced to the stage of sophistication where we do not believe that the universe was made for the sole benefit of mankind here on Earth.

We know more about the Earth's surface and its interior than we do about any other planet; consequently, in our study of the solar system we can take advantage of our knowledge of a particular planet. This knowledge will add greatly to our understanding of all the planets and perhaps help us understand how they were all formed.

Most of the ideas about the Earth's interior are based on information gained from the study of earthquake waves. As these waves travel through the Earth, their speed and direction are changed according to the type of wave and the material through which they are traveling.

These studies show that it is convenient to consider the Earth as four concentric spheres: the inner core, the outer core, the mantle, and the crust (Figure 4-3). The core is composed principally of iron and nickel; the inner core is believed to be solid, the outer core molten. Above the core is the mantle, composed chiefly of the denser rock-forming substances. Atop the mantle is the crust, the rocky surface of the Earth. It will be seen in Chapter 7 that knowledge of the Earth's iron-nickel core has influenced our thinking about the formation of the solar system.

The Earth's atmosphere has influenced the surface of the Earth a great deal. The atmosphere carries the water vapor which forms the rains, and the rains have been the chief agent in erosion and transportation of material from one location on the crust to another. The crust has responded to this relocation of material by sinking where the material was deposited and rising in that location from which the material was taken, changing the patterns of erosion.

The atmosphere actually contains very little water vapor. Nitrogen

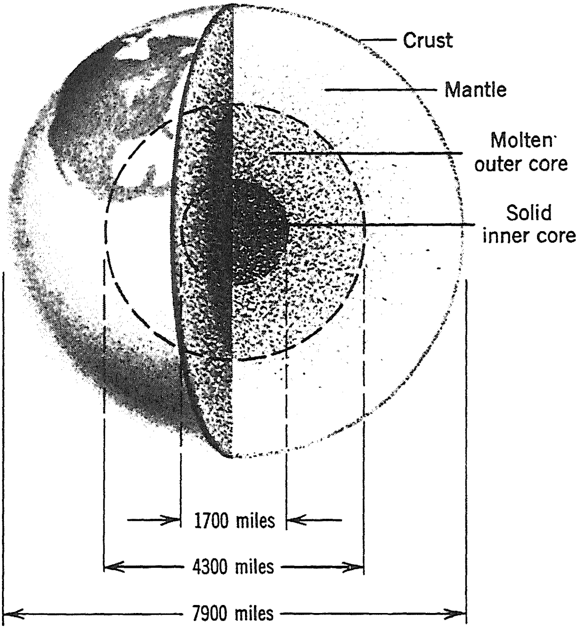


Figure 4-3 A cross section of the Earth.

is the dominant gas in the atmosphere (78%), oxygen is next (21%), with only 0.03% carbon dioxide. The rest of the gases comprise the very small percentage that remains.

As valuable as the atmosphere is, the astronomer is anxious to transport telescopes above and beyond it. Earthbound telescopes permit the astronomer to examine the light and radio waves of celestial objects only after they have passed through the atmosphere and its disturbing influence. Severe restrictions are placed upon the terrestrial astronomer because the atmosphere acts like a filter, and permits only certain regions of the electromagnetic spectrum to pass through, namely the visible region and large regions in the radio wavelengths. It blocks out all the other radiations such as X-rays, and much of the ultraviolet and infrared, which contain a great deal of information about the universe. Consequently space explorations have made major contributions by observing the sun and the celestial sphere from above the Earth's atmosphere in X-ray and ultraviolet radiation. These observations (like most completely new observations) tell us not only what we have been missing but also that we have much yet to learn.

The light that does reach our terrestrial telescopes has some alterations imposed upon it by the atmosphere. The atmosphere acts somewhat like a lens, for by refraction it changes the direction of light passing through it (unless that light is traveling vertically). Thus the position of each object in the sky is shifted slightly toward the *zenith* (the point directly overhead) from its true position. The shift is zero for stars in the zenith and increases for stars as they approach the horizon. At the horizon the shift amounts to about 30 minutes of arc, or nearly the angular diameter of the sun and the moon. Therefore when you see the sun (and the moon) sitting on the horizon it has actually already set and is just below the horizon.

Not only does refraction cause a star's position to change in this systematic manner, but also the differential refraction caused by convection currents in the atmosphere (similar to those visible above a hot road) causes a star's image to move about slightly and erratically. To the eye this is twinkling, which may be poetic and romantic. The astronomer, however, calls it "bad seeing," for the unpredictable movement of a star's image can only result in a blurred photograph, or a decrease in the amount of light passing through the slit of a spectrograph.

It is the atmosphere that gives rise to the blue sky which outshines the stars in the daytime. Sunlight that enters the atmosphere is intercepted not only by the gases that give rise to absorption lines but also by tiny dust particles. These dust particles, as well as the gas molecules, act somewhat like reflectors in that they can change the direction of light's travel. The process, which is different from reflection, is called *scattering* and is selective in that molecules and particles scatter the blue light (the shorter wavelengths) more effectively than they scatter red light. Since the blue is scattered more thoroughly, our sky is blue for it consists of this scattered light; the setting sun is red and the harvest moon orange because we see them after their light has had much of the blue taken out. When the sun and the moon are on the horizon their light travels through a longer path in the atmosphere, with the result that more of the blue is scattered out of the direct beam of light.

From above the Earth's atmosphere, rockets and space probes have made observations and measurements of features not dreamt of before. Two of the first discoveries of the space age are the Van Allen radiation belts and the extended magnetic field of the Earth. Radiation counters aboard satellites and space probes sent to both Venus and Mars have revealed that the Earth is surrounded by a large magnetic field, a portion of which traps high-energy electrons (Figure 4-4).

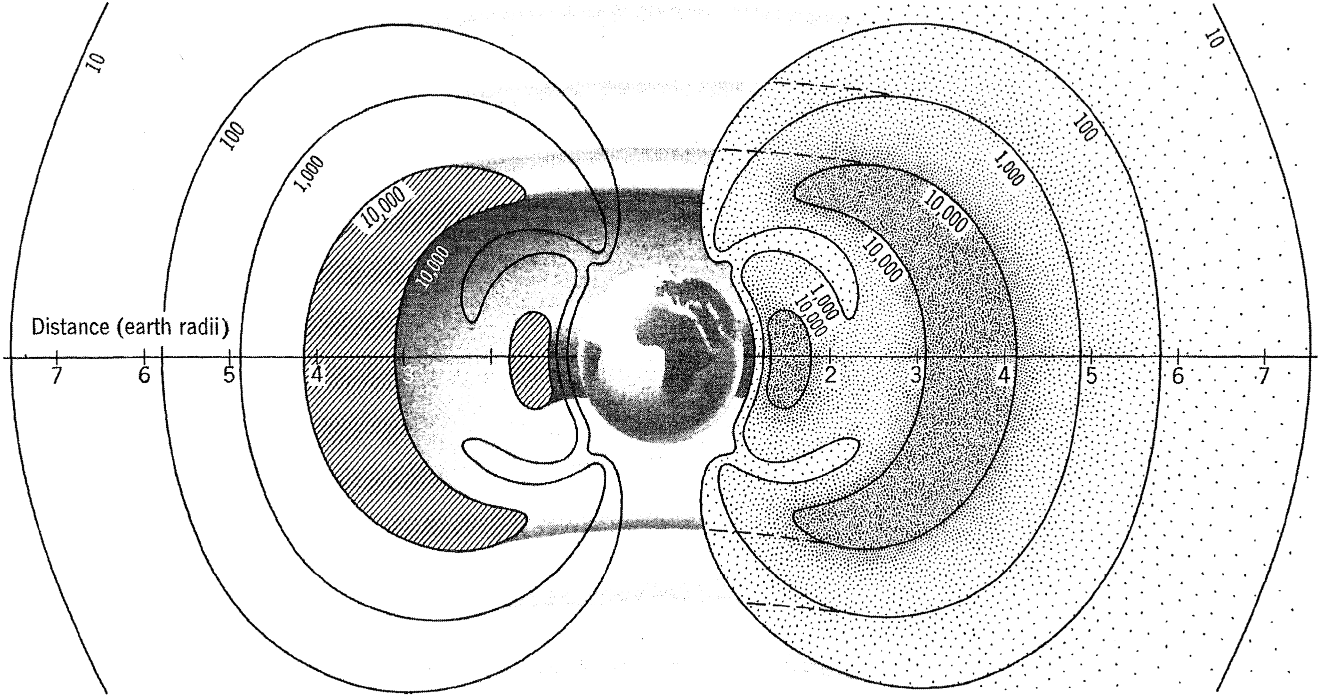


Figure 4-4 Van Allen radiation belts. (Reprinted by permission from *Scientific American*)

These electrons travel in corkscrew paths from the region of one pole to the other, along the magnetic lines of force. In the polar regions, these electrons dip into the upper atmosphere, causing it to glow in much the same manner that a gaseous tube glows. This glow is observed as the *aurora,* both the aurora borealis (northern lights) and the aurora australis (southern lights).

The magnetic field of the Earth, however, is not free from external disturbing influences. Space probes have detected the existence of what is called the *solar wind,* a stream of charged particles, mostly electrons and protons, emitted by the sun and flowing past the Earth. Since these particles are charged, they carry a magnetic field with them. This magnetic field reacts with the magnetic field of the Earth in a way which appears to "blow" the magnetic field away from the Earth.

On the sunward side, the Earth's magnetic field extends to a distance of about 10 Earth's radii. Just beyond 10 Earth's radii, a region of magnetic turbulence has been observed, with an outer boundary consisting of a shock wave (Figure 4-5). The shock wave is the encounter between the particles of the solar wind and the Earth's magnetic field. It resembles a shock wave set up by a rock projecting above the surface of a river, or the wave set up by the bow of a ship as it cuts through the water. Outside this shock wave, the magnetic field is that of the solar system.

On the "leeward" (nighttime) side of the Earth the magnetic field

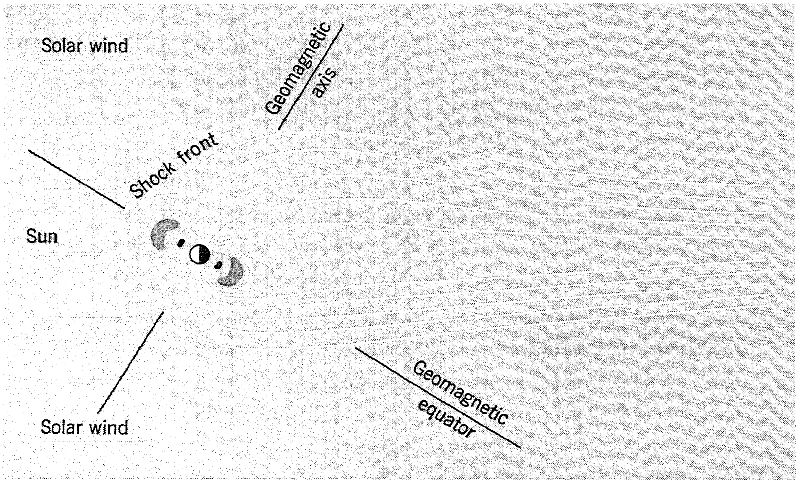


Figure 4-5 The Earth's magnetic tail caused by the solar wind.

forms a "magnetic tail." The Earth's magnetic field has been detected in this direction as far away as 25 Earth's radii, and it appears that the tail is about 40 Earth's radii in diameter. The magnetic lines of force in the northern part of the magnetic tail point toward the Earth, that is, the north end of a compass needle if placed there would point to the Earth. The magnetic lines of force in the southern part of the tail point away from the Earth. The boundary between these two regions of oppositely directed magnetic fields within the tail shows very little—if any—magnetic field and is called the *neutral surface.* The similarity between the shape of the Earth's magnetic tail, with its encircling region of magnetic turbulence, and the shape of a comet's tail is striking.

Mars

Mars, the "red planet of mystery," is the fourth planet we investigate as we go farther away from the sun. We are able to study Mars better than we can any of the other planets because of its proximity to us, and because when it is closest to the Earth we see its fully lighted face.

Mars, 1.5 A.U. from the sun, rotates on its axis in 24 hours, 37 minutes, and revolves about the sun once every 687 days. Its diameter, 4,200 miles, is about ½ that of the earth; its mass is 0.39 times that of the Earth. It has an atmosphere, but one which is not so thick that we cannot see through it. It has permanent surface markings and polar caps. The markings change color during the Martian year and we can see cloud formations that move across the surface of the planet. Its temperature presents extremes. Measurements of the temperature at the equator at local Martian time are: at 7 A.M. —52°F, at 12:30 P.M. +63°F, at 2 P.M. +46°F; and it is expected that the coldest, reached at about 5 A.M., is about —66°F. This amounts to a daily change in temperature of about 130°F.

The principal reason for the great variation of temperature during the Martian day is the comparative lack of atmosphere. The most accurate measurements of the Martian atmosphere were made on July 15, 1965, when the space probe, Mariner IV, passed behind Mars. During passage behind the planet, the radio aboard Mariner IV transmitted a prearranged signal that passed through the Martian atmosphere. The space probe went behind the planet on the daytime side and emerged on the nighttime side, so the signal passed from top to bottom through both the sunlit and the dark atmosphere.

It was observed that the sunlit atmosphere has an extensive layer of ionized gas, but the nighttime atmosphere of Mars has very little, if any, ionization. It was also determined during the fly-by that the strength of the magnetic field of Mars is less than one-thousandth of the strength of the Earth's magnetic field. Therefore, Mars does not have an extensive radiation belt similar to the Van Allen radiation belts of the Earth.

Other results of this spectacular experiment show that the pressure at the bottom of the Martian atmosphere is even less than expected. For comparison, the atmospheric pressure here on the surface of the Earth is close to 1,000 millibars (a pressure of 1,013 millibars equals 14.7 pounds per inch squared); the atmospheric pressure on the surface of Mars is between 4 and 7 millibars, the uncertainty arising because of uncertainties in the chemical composition of the atmosphere.

There is positive evidence obtained by spectrographic work that the Martian atmosphere contains both carbon dioxide and water vapor, but very little, if any, oxygen. If all of the water in the atmosphere of Mars were to condense, it would form a layer 0.04 in. deep over an ideally spherical Mars. No fishing up there!

Evidence of the existence of water vapor is not only spectroscopic, but direct. Mars has polar caps similar to those on Earth, but much less extensive. These caps reflect light in much the same manner as frost. One cap recedes when its hemisphere is in summer, while at the same time the other polar cap grows during the winter. A haze or mist forms over a pole during its late autumn, and stays there until winter has run its course. As spring comes the haze lifts and reveals the snow underneath. The cap then decreases in size until by the end of summer there may be no polar cap at all. Since the temperature of Mars is less than that of the Earth, it must be concluded that the polar caps cannot be over a few inches thick. If they were as thick as the Earth's polar caps they would certainly not melt completely during the summer.

As the south polar cap melts each Martian spring, the edge does not retreat uniformly. Some parts of it melt more slowly than others, indicating that Martian topography is not regular. In fact, there is one portion that remains white much longer and even becomes detached from the rest of the polar cap (Figure 4-6). We are led to assume that this area, called the Mountains of Mitchell, retains the snow longer because it is higher (or lower).

As the south polar cap of Mars melts, a faint band of darkening

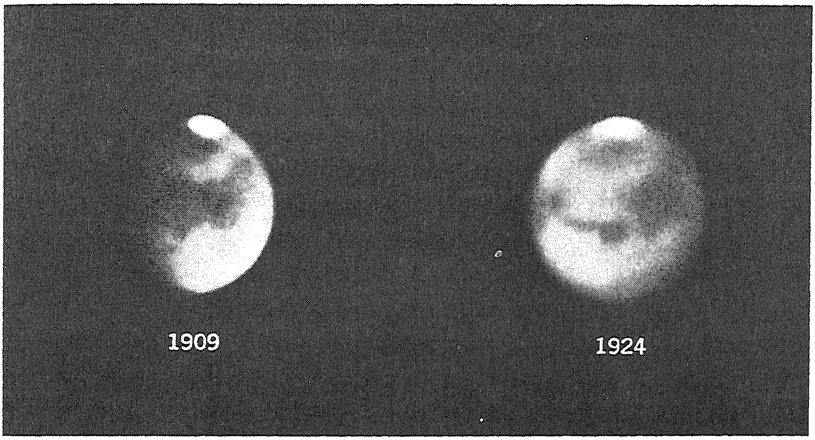


Figure 4-6 Two photographs of Mars taken in 1909 (left) and 1924 (right), showing the same small detached portion of the south polar cap. This patch of snow is called the Mountains of Mitchell. Both photographs were taken on the same Martian date, June 3. Since the telescope turns the image upside down, astronomers are accustomed to observing astronomical objects with the south at the top. (Lowell Observatory)

develops immediately to the north around its edge. With the advance of the season this darkening travels as a wave toward the equator and the permanent markings become even darker. The wave does not stop at the equator as one might expect, but continues into the northern hemisphere. Whether it is water in a liquid state or water vapor has not yet been determined. The fact that it travels at a fairly uniform rate (22 miles per day) and continues on into the northern hemisphere would indicate that it is water vapor carried by the atmosphere. There are a few portions of the Martian surface that become much darker than the others when the wave passes. It has been suggested that these darker areas may result from the presence of water in the liquid state.

The permanent dark markings, of course, have commanded much attention. Some of them appear to be straight lines. It is these apparent straight lines, first seen by Schiaparelli in 1877, that have given rise to the idea of artificial waterways on Mars. Not only their apparent straightness but also their apparent interconnections has convinced some that these dark lines are *canals* built by intelligent beings. However, from what we now know, it appears quite improbable that intelligent beings could exist on Mars.

Other permanent markings are widespread, dark areas which, because of their appearance and permanency, were named *maria* (sing. mare). Maria means "seas," but we know now that they are not seas, or oceans, or even swamps. It is principally the maria and the canals that appear to change color during the Martian year. For most of the seasons they are bluish green or bluish grey, but during the winter they turn a rather brownish color.

Attractive as these results seem, it must be remembered that observations of color are more difficult than many other types of observation. Most of the surface of Mars has been described as having a brick red or orange color. Investigations by two astronomers working independently and using different methods have shown that the blue-green markings mentioned above are not actually blue-green. They appear blue-green only by contrast with the brick color of the rest of the planet. Apparently, casual observation may exaggerate the strength of a tint which, though different, does not constitute the dominant color of the darker areas.

But the fact remains that these areas are different from the rest of the planet. Are they regions of plant growth? Are they regions of different soil? Is it possible that certain chemicals could account for the different color of these areas as well as for their seasonal change of color?

It may well be that the canals result from the manner in which the craters of Mars are scattered about the surface. Mariner IV transmitted back to Earth 22 photographs of the Martian surface. At least in its gross features, the Martian surface is not unlike that of the moon. It is scarred with craters (Figure 4-7), and the number of craters per unit area is about the same as the moon; their depth and size seem to be about the same, some of them have peaks in the middle, and some are very big and others are small. Both photographs shown in Figure 4-7 straddle a boundary between a dark marking and an unmarked surface. Yet the marking does not appear on either photograph. In photograph number 9, an orange filter was used; photograph number 11 was made through a green filter.

But if the Mariner IV fly-by settled the question of the basic surface features, it did not settle the question of life on Mars—nor was it expected to. The photographs simply do not contain enough detail to make such a critical analysis of the surface. Should there be life on Mars, evidence to support that fact will have to be acquired during other space probes.

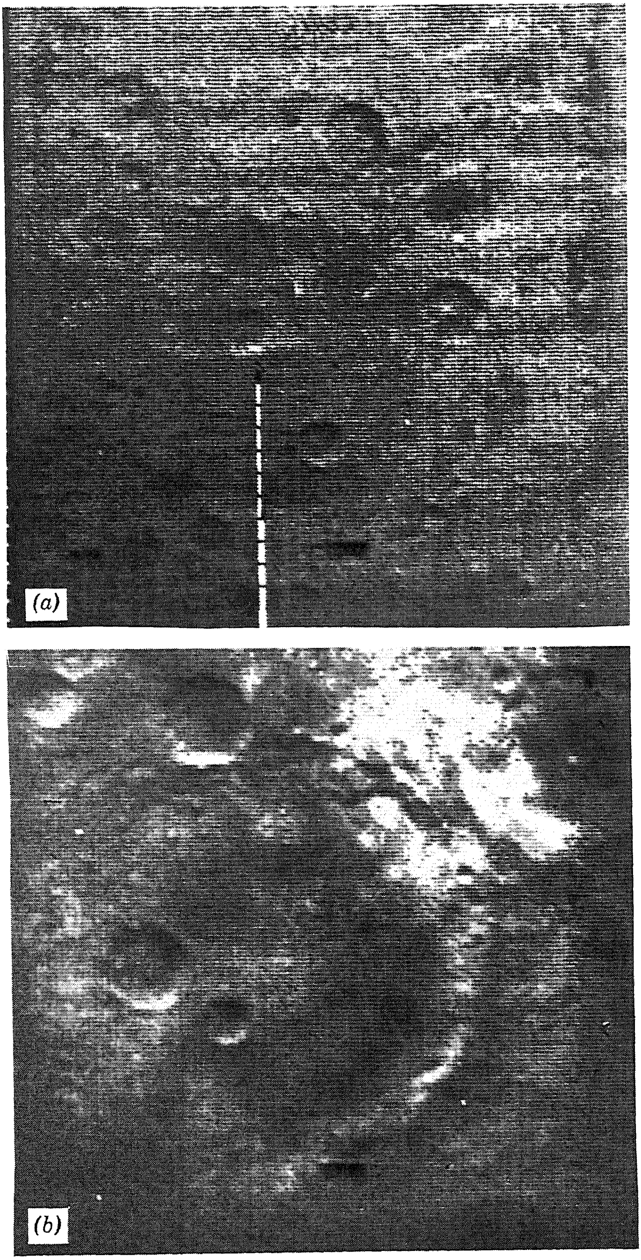


Figure 4-7 (*a*) Mariner IV photograph No. 9 of the Martian surface. At this time the camera was 8,100 miles from the region it photographed. (*b*) Mariner IV photograph No. 11. These remarkable photographs clearly indicate that the Martian surface is crater-marked.

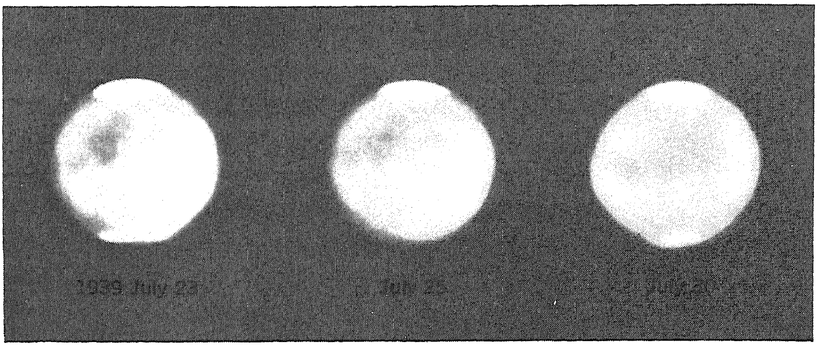


Figure 4-8 Three photographs of Mars taken on three different dates in 1939—July 23, July 25, and July 30—to show the formation, disappearance, and reformation of clouds over the north polar region in just a few days. (Lowell Observatory)

Various types of observations indicate that this basic surface, with its brick color, might be composed largely of a mineral called limonite, which contains a good deal of iron oxide. On Mars, it appears to be in a powdery form. Could this account for the lack of oxygen in the atmosphere? Has most of it been trapped by a rusting away of iron in the soil?

A surface of powder would help to account for the large yellow clouds that appear from time to time on Mars. These yellow clouds at times cover a very large portion of the planet's surface and are attributed to large dust storms. They seem to lie low in the atmosphere as dust storms would.

They are not the only clouds, however, on this planet of mystery. The clouds over the polar regions have already been mentioned. In Figure 4-8 the south polar cap, at the top of the picture, is quite large, but the size of the north polar cap varies. On July 23 a white cloud appears which is to disappear 48 hours later. Five days later, July 30, another cloud appears.

White clouds appear in other regions of the planet as well. The clouds may form about noon and grow larger during the afternoon until they seem to loom over the edge of the planet at twilight. Whitish areas are sometimes visible during a Martian morning and disappear as the temperature increases during the day. Since these are not as high as clouds, they have been explained as formations of frost or fog which dissipate under the noon sun.

There is a third type of cloud, called the "blue cloud," though to the eye it does not appear blue. Such a cloud is most pronounced on photographic plates and its blueness has been deduced from the fact that it shows up well on plates sensitive to blue light, but not on red-sensitive plates. (The reverse is true of the yellow clouds, which show up better on red-sensitive plates.) These blue clouds, which lie higher in the Martian atmosphere than the white clouds, may be composed of ice crystals like the cirrus clouds of the Earth.

Surface features of Mars are best studied with photographic plates and filters sensitive to red rather than blue light. Blue light is scattered more than red by dust particles and air molecules; consequently a photograph of Mars in blue light shows for the most part the atmosphere from which the sun's blue light was scattered.

Strangely enough, however, there are times when the surface of Mars can be photographed in blue light. That which causes the blue to be scattered and normally prevents us from photographing the surface in blue light evidently dissipates during these periods of "blue clearing." Then, blue light strikes Mars' surface and is reflected back out of its atmosphere with minimal scattering, though for what reasons we do not yet know.

Jupiter

The fifth and largest planet is Jupiter. Its mass is 318 times that of the Earth and its diameter is 86,700 miles, over 10 times the Earth's diameter. It is about 5.2 A.U. from the sun. Because of its enormous distance from the source of heat its temperature is quite low, approximately —200°F. There are markings which enable us to determine its period of rotation near the equator at 9 hours, 50 minutes. The period of rotation increases at latitudes nearer the poles, to about 9 hours, 56 minutes. This tells us not only that the surface we see is not solid but also that someone at a latitude of 45° north or south of the equator would need to have a clock whose rate was different from that of a clock on the equator, as well as a different calendar.

Jupiter is not only the largest planet but also rotates on its axis in a shorter period than any other planet. The material at the equator whirls about the axis at about 30,000 miles per hour. (The Earth's equator whirls at about 1,000 miles per hour.) As a result, Jupiter is flattened at the poles to the extent of presenting a slightly elliptical rather than a circular disk.

The markings on Jupiter consist of cloud formations: not clouds like those on the planets nearer to the sun, but clouds composed of methane and ammonia gas as well as an abundance of helium and molecular hydrogen. Methane gas is frequently used by city gas companies, so the gas that is burned in your stove may be methane. Ammonia is a poisonous gas with an odor which once smelled will not be forgotten.

The clouds form belts or bands about Jupiter because their period of rotation depends on their latitude. The clouds north and south of the equator rotate more slowly than those on the equator, forming streaks such as one would get when mixing melted chocolate in the batter for marble cake. The bands change over the years, and that change is sometimes fairly rapid. Photographs taken at New Mexico State University Observatory reveal that the changing character of the belts can be dramatic (Figure 4-9). Between October 1964 and February 1966 the Equatorial Zone changed from one of the darker features on the cloud surface to one of the lightest features. In 1964 the South Tropical Belt—containing the Great Red Spot—was light, but in February 1966 it contained a dark streak.

There is one notable feature in these clouds that does remain fairly constant. It is a large spot, often a dull brick red in color, which is 7,000 miles wide and 30,000 miles long. It is shown in Figure 4-9. The Great Red Spot was first seen about 300 years ago and has not changed appreciably since, although its brightness does vary as well as its color. It has been observed to move through the surrounding clouds up to 8,000 miles in one (terrestial) day, which indicates that it probably is not attached to any solid surface below the clouds. The true nature of the Great Red Spot is not known.

White spots, too, appear from time to time; they also move in relation to the clouds but are rather short-lived. They generally smear out into a belt or band.

The motions of the clouds indicate that they are not bound by any underlying solid surface. This gives us a clue in our study of Jupiter's interior. Although Jupiter has a mass 318 times the mass of the Earth, its volume is 1,312 times that of the Earth. Consequently the material of the Earth, volume for volume, must have a mass that is about 4 times as great as that of Jupiter. This can be expressed in terms of *density*.

The density of a substance is its mass per unit volume. The density of water may be taken as 1 and all other materials can then be compared to water. The average density of the Earth is 5.5 and the density of Jupiter is only 1.3. Consequently, volume for volume, the

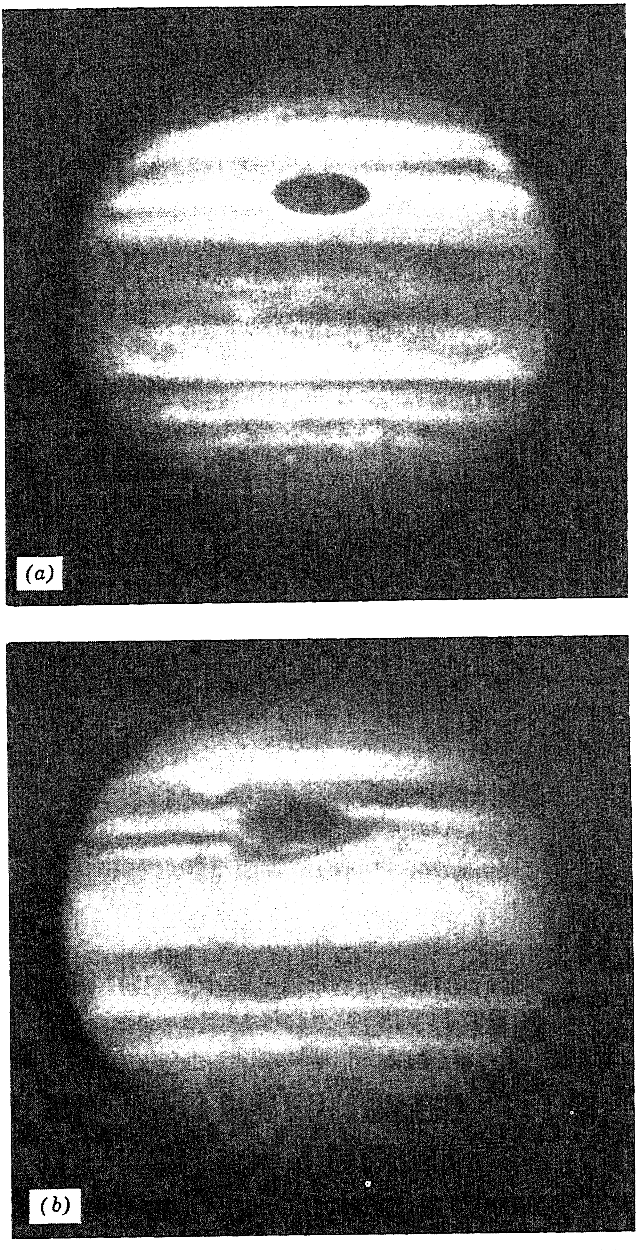


Figure 4-9 (*a*) Photograph of Jupiter taken on October 23, 1964 shows the Tropical Belt dark. The Great Red Spot is obvious. (*b*) This photograph was taken on February 21, 1966 and the Tropical Belt is bright. Even the belt containing the spot has changed in the intervening months between these two photographs. (New Mexico State University Observatory)

Earth has a mass that is 5.5 times as great as water, whereas Jupiter is only 1.3 times as massive. This low density indicates that Jupiter is about three-fourths hydrogen.

According to one theory, the atmosphere of Jupiter may be less than 500 miles thick, and may rest upon and merge into a sea of liquid hydrogen. With increased depth, the corresponding increase in pressure must compress the hydrogen into a solid. In the very center may be a rocky core not much larger than the Earth.

This concept of Jupiter's structure is supported by the oblateness of its surface. The diameter from pole to pole is 83,200 miles, and the equatorial diameter is 88,700 miles. This is about 5,500 miles out of about 85,000, considerably more than the Earth's oblateness of 27 miles out of about 8,000. Jupiter's greater oblateness can be accounted for not only by its more rapid rotation but also by the theory mentioned above that Jupiter must be much denser at the center than near the surface.

Besides reflecting the sun's light, Jupiter emits its own radiation in the form of radio waves. This discovery, made in 1955 by K. L. Franklin and B. Burke, then with the Carnegie Institute, opened up a whole new field of planetary radio astronomy, and gave us a new tool with which to probe the secrets of our solar system. Jupiter has been studied intensively since 1955, not only with radio telescopes, but with increased optical interest. The attempt is to correlate as many observations as possible.

There are two basically different radio emissions from Jupiter: the *microwaves*, with a wavelength between 3 and 68 cm, and the *decameter waves* with wavelengths between 10 and 50 meters. The decameter range is more often referred to by its frequency between 5 and 25 megacycles per second. These will be considered one at a time.

The microwaves are difficult to observe and there is some question as to their nature. For example, by using the radio energies in this range to compute the temperature of Jupiter, discordant calculations result. At a wavelength of 3 cm, the surface temperature is calculated to be —210°F, in close agreement with optical studies. But as longer and longer wavelengths are used, the temperatures become more appropriate for a star than for a planet. Therefore it is necessary to assume that this radio energy is emitted by nonthermal sources, perhaps by electrical discharges, or more regular electron motion. Through the use of radio interferometers to more nearly pinpoint the source of the microwaves, it has become clear that they originate not on the surface but in radiation belts similar to the Van Allen radiation belts about the Earth.

The radiation belts about Jupiter are evidence that that planet also has a magnetic field, and one much stronger than the Earth's. The radiation belts about Jupiter extend out to a distance of more than 130,000 miles, far enough to include the innermost satellite, Number V. (It was the fifth satellite to be discovered and happens to be closer to the planet than the four which Galileo discovered.)

The decameter waves which arrive are much more energetic than the microwaves and consequently are more easily studied. These waves may appear in bursts lasting only a few seconds, or in "noise storms" lasting many hours. One source of difficulty in studying the decameter waves is that the Earth's atmosphere makes them scintillate or twinkle just as stars twinkle at night. Observations made by two radio telescopes placed 40 miles apart were quite different. Each radio telescope may pick up "bursty" type signals, but there is no correlation between the bursts received from each of the telescopes. Again the need to place telescopes above the Earth's atmosphere becomes apparent.

It has become clear, however, that the decameter waves are associated more closely with the surface than the microwaves. In fact, the decameter waves exhibit a periodicity close to the period of rotation of the planet itself. They would appear to originate from a definite region on the planet.

Since the period of rotation of Jupiter depends upon latitude, two periods of rotation were given in the past: the first is called system I and relates to the equatorial rotation. Its period is given as 9 hours, 50 minutes, 30.003 seconds. The second is system II and is associated with the rest of the planet; its period is 9 hours, 55 minutes, 40.632 seconds. The period of rotation observed by decameter radio waves is 9 hours, 55 minutes, 29.37 seconds and is called system III.

Recently a noticeable correlation has been established between the revolution of Io, the second satellite from Jupiter and thus the innermost of the Galilean satellites, and the decameter radio bursts. It seems that the radio bursts occur more frequently when the satellite Io is above the sunset terminator of Jupiter, that is, at its westernmost position; or when Io is approaching its easternmost position as it revolves about Jupiter in its period of 1 day, 6 hours, 18 minutes. It may be that Io causes ionization in the extreme upper atmosphere of Jupiter and this ionization results in electric discharges of one kind or another.

Saturn

Saturn, 9.5 A.U. from the sun, is perhaps the most spectacular astronomical sight in a telescope, because of the fantastic rings which encircle it, and which have long been used by cartoon artists to set the scene out in space. Saturn revolves about the sun in 29½ years. Its oblateness is more extreme than that of Jupiter; its polar diameter is 67,900 miles and its equatorial diameter is 75,100 miles. The period of rotation at the equator of Saturn is 10 hours, 14 minutes, but with increasing latitude its period of rotation increases even more rapidly than with Jupiter. The density of Saturn is more extreme than that of Jupiter; at 0.68 it is even less than that of water—even less than that of butter! Its mass, however, is 95 times that of the Earth.

Saturn's atmosphere is similar to that of Jupiter except that since it is colder on Saturn (—234°F) some of the ammonia has crystallized out of the gaseous state. Therefore Saturn's atmosphere has in the gaseous state relatively more methane and less ammonia. (The solid particles, of course, do not add any dark lines to its spectrum.)

Saturn too has belts or bands somewhat like Jupiter's, though they are much less pronounced. These also change from time to time. Occasionally, white spots appear which merge into a white belt and then disperse into the atmosphere.

The multiple rings are what make Saturn unique in our solar system They have a maximum diameter of 170,000 miles, are concentric in a single plane, cannot be much thicker than 10 miles and may be thinner. Looking through a large telescope one can see gaps between them. The outer gap, called Cassini's division, is seen in Figure 4-10. There is also an inner division.

The composition of the rings was a major puzzle for astronomers until the advent of the spectrograph. Around the turn of the century an astronomer named Keeler placed the image of the planet formed by his telescope across the slit of a spectrograph in order to find the rate of rotation of the rings. If the rings are solid then the outer part should rotate faster since it has farther to go in one rotation. If they are not solid, however, then the inner portion should have a higher velocity than the outer. This is derived from the same principle (discussed in Chapter 5) that explains why Mercury revolves

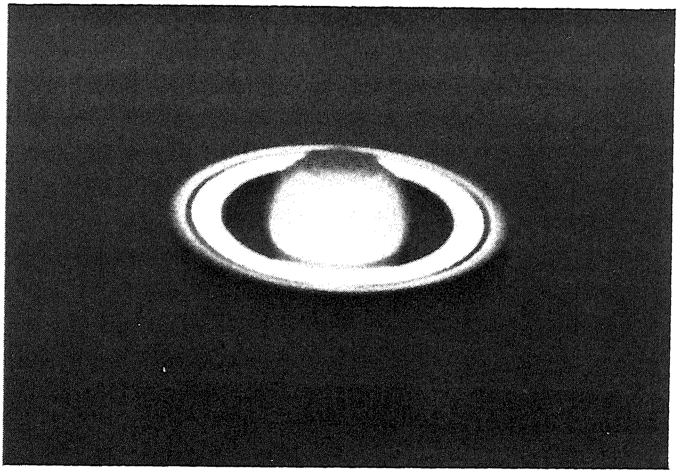


Figure 4-10 A photograph of Saturn with its ring system showing Cassini's division. (Lick Observatory)

about the sun faster than the other planets, and why Pluto, the planet most distant from the sun, revolves the slowest in its orbit.*

Keeler determined the velocities of approach and recession of the limbs of both Saturn and its rings by the Doppler principle. This critical test revealed that the inner portion of the rings rotates more rapidly than the outer portions and therefore the rings are not solid.

But if the rings are not solid, what are they? So far as we know they yield no spectral lines of their own and therefore cannot be gaseous. They reflect the sun's light in such a manner that we are led to conclude that they are composed of small rock particles or pebbles coated with ice, each of which has its own orbit about Saturn. As a further confirmation of this concept, it should be noted that as Saturn moves against the background of stars it occasionally passes in front of a bright star. When this happens the bright star, though dimmed, can be seen through the rings.

Viewed from the Earth the rings present a different aspect depending on how they are tilted with respect to our line of sight. If the rings were parallel to and in the plane of the Earth's orbit, we would always see them edge on and since they are so thin they would have escaped notice for a much longer time. But they are tilted from and do not lie in the plane of the Earth's orbit; consequently at of Jupiter and this ionization results in electric discharges of one kind or another.

---

* The orbital velocity of Mercury is 29.7 miles per second whereas Pluto's orbital velocity is only 3.0 miles per second.

Saturn

Saturn, 9.5 A.U. from the sun, is perhaps the most spectacular astronomical sight in a telescope, because of the fantastic rings which encircle it, and which have long been used by cartoon artists to set the scene out in space. Saturn revolves about the sun in 29½ years. Its oblateness is more extreme than that of Jupiter; its polar diameter is 67,900 miles and its equatorial diameter is 75,100 miles. The period of rotation at the equator of Saturn is 10 hours, 14 minutes, but with increasing latitude its period of rotation increases even more rapidly than with Jupiter. The density of Saturn is more extreme than that of Jupiter; at 0.68 it is even less than that of water—even less than that of butter! Its mass, however, is 95 times that of the Earth.

Saturn's atmosphere is similar to that of Jupiter except that since it is colder on Saturn (—234°F) some of the ammonia has crystallized out of the gaseous state. Therefore Saturn's atmosphere has in the gaseous state relatively more methane and less ammonia. (The solid particles, of course, do not add any dark lines to its spectrum.)

Saturn too has belts or bands somewhat like Jupiter's, though they are much less pronounced. These also change from time to time. Occasionally, white spots appear which merge into a white belt and then disperse into the atmosphere.

The multiple rings are what make Saturn unique in our solar system They have a maximum diameter of 170,000 miles, are concentric in a single plane, cannot be much thicker than 10 miles and may be thinner. Looking through a large telescope one can see gaps between them. The outer gap, called Cassini's division, is seen in Figure 4-10. There is also an inner division.

The composition of the rings was a major puzzle for astronomers until the advent of the spectrograph. Around the turn of the century an astronomer named Keeler placed the image of the planet formed by his telescope across the slit of a spectrograph in order to find the rate of rotation of the rings. If the rings are solid then the outer part should rotate faster since it has farther to go in one rotation. If they are not solid, however, then the inner portion should have a higher velocity than the outer. This is derived from the same principle (discussed in Chapter 5) that explains why Mercury revolves

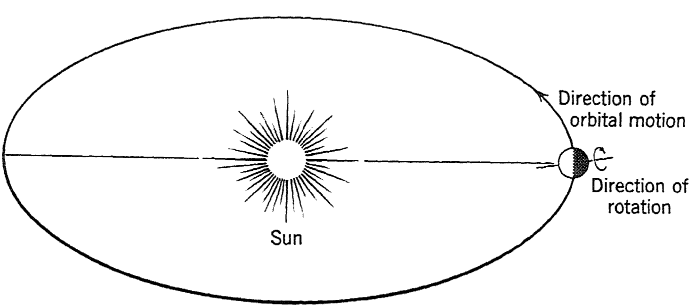


Figure 4-11 The axis of rotation of Uranus is tipped by 98° from the perpendicular to the plane of its orbit. To compare with the Earth see Figure 5-7.

The most unusual thing about Uranus is the inclination of its axis of rotation. The Earth's axis of rotation is not perpendicular to the plane of its orbit but is tipped 23½° from this perpendicular. The axis of rotation of Uranus, however, is tipped 98° from the perpendicular to its orbit. Or, to put it another way, its axis is tipped by only 82°, if we consider it as rotating in retrograde motion (Figure 4-11). Just how Uranus got this way is a puzzle which must be answered by any satisfactory theory of the origin of the solar system.

After Uranus was discovered the astronomers observed its position carefully over a period of years in order to determine its orbit, that is, where it goes and how fast. Newton's laws of motion and his law of gravity (see Chapter 5) describe the motion of the planets about the sun very well and can be used to predict the position of a planet or comet once its orbit has been determined, but Uranus did not move as expected. Two astronomers, Adams and Leverrier, noted this discrepancy and independently both concluded that Uranus was not following its prescribed path because an assumed planet, still more distant from the sun, was disturbing its motion. By Newton's law of gravity they both calculated where such a planet should be.

Leverrier made contact with an observatory that had been preparing a star map. By comparing the position of the predicted planet on the star map with the actual sky, the new planet was seen within a matter of minutes. This was a tremendous triumph for Newton's law of gravity; not only could it account for the motions of the planets but it enabled the astronomer to detect heretofore unknown planets by their gravitational effect on already known planets, *and* to determine their positions!

times we see them edge on and at other times we see them at an oblique angle.

Uranus

The planets discussed thus far are bright enough to be seen with the naked eye and consequently have been known since prehistoric times. Uranus, the first planet discovered in recorded history, is the next to be seen as we go out from the sun. One of the greatest astronomical observers of all time, William Herschel, discovered it in 1781. The story of the discovery indicates the value of better instruments and experience in observing.

Herschel built his own telescopes as an amateur and devoted his life to observing. Those were days when but few of the spectacles of the sky were known, and it was Herschel who discovered many of the objects that today are being investigated so that we may better understand how the universe is put together and why it stays together—if it is staying together.

Herschel spent much of his time scanning the skies looking for objects that were not stellar in appearance, objects such as the large nebulae we discuss in a later chapter. One of these objects presented a disk-like structure which made it appear different not only from the stars but from the nebulae as well. On magnifying it further Herschel thought it was a comet, but later it was recognized as the first discovered planet. Moreover, it had already been seen by one other astronomer who at that very time was preparing a star map of the sky and had in fact recorded the planet's position twelve times. But he had thought of it only as a star, partly because his telescope was not as large as Herschel's and partly because it was stars with which he was concerned.

Uranus is about 19.2 A.U. from the sun, with a mass 14 times as large as that of the Earth. It revolves about the sun in 84 years and rotates on its axis in 10 hours, 49 minutes. Its diameter is about 30,000 miles. Because of its great distance from the sun its surface temperature, at —300°F, is even less than that of Saturn. Its atmosphere is similar to that of Jupiter and Saturn, except that even more of the ammonia has crystallized, leaving an atmosphere with an even higher percentage of methane gas. Recently molecular hydrogen has been discovered on Uranus as well as on Jupiter.

Uranus is so far from the Earth that any surface markings it may have are very difficult to see, so its period of rotation had to be determined by the spectrograph.

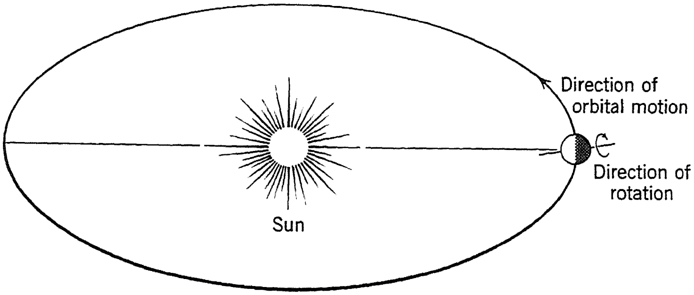


Figure 4-11 The axis of rotation of Uranus is tipped by 98° from the perpendicular to the plane of its orbit. To compare with the Earth see Figure 5-7.

The most unusual thing about Uranus is the inclination of its axis of rotation. The Earth's axis of rotation is not perpendicular to the plane of its orbit but is tipped 23½° from this perpendicular. The axis of rotation of Uranus, however, is tipped 98° from the perpendicular to its orbit. Or, to put it another way, its axis is tipped by only 82°, if we consider it as rotating in retrograde motion (Figure 4-11). Just how Uranus got this way is a puzzle which must be answered by any satisfactory theory of the origin of the solar system.

After Uranus was discovered the astronomers observed its position carefully over a period of years in order to determine its orbit, that is, where it goes and how fast. Newton's laws of motion and his law of gravity (see Chapter 5) describe the motion of the planets about the sun very well and can be used to predict the position of a planet or comet once its orbit has been determined, but Uranus did not move as expected. Two astronomers, Adams and Leverrier, noted this discrepancy and independently both concluded that Uranus was not following its prescribed path because an assumed planet, still more distant from the sun, was disturbing its motion. By Newton's law of gravity they both calculated where such a planet should be.

Leverrier made contact with an observatory that had been preparing a star map. By comparing the position of the predicted planet on the star map with the actual sky, the new planet was seen within a matter of minutes. This was a tremendous triumph for Newton's law of gravity; not only could it account for the motions of the planets but it enabled the astronomer to detect heretofore unknown planets by their gravitational effect on already known planets, *and* to determine their positions!

Neptune

The new planet was named Neptune. It is about 30 A.U. from the sun, has an equatorial diameter of 28,000 miles (it is not as oblate as Jupiter), and revolves about the sun in 165 years. The mass of Neptune is 17 times that of the Earth. It has a period of rotation a little longer than 15 hours and a surface temperature of about —330°F. Aside from knowing that Neptune has stronger bands of methane in its spectrum than any other planet, astronomers know little about the planet.

Uranus was discovered in 1781 and Neptune was discovered in 1846. These planets travel so slowly (their average orbital velocities are respectively 4.2 and 3.4 miles per second), that it takes a long time to determine an accurate orbit for them. By the early 1900's Lowell, the founder of the Lowell Observatory in Flagstaff, Arizona, had reobserved Uranus and concluded that the discovery of Neptune could not account for all the discrepancies in the orbit of Uranus. He therefore predicted that there must be another planet even farther away than Neptune. The new planet was discovered by Clyde Tombaugh in 1930 at the observatory in Flagstaff and named Pluto.

Pluto

Pluto travels around the sun in 248 years and with an average velocity of 3.0 miles per second. It is about 40 A.U. from the sun and its surface temperature seems to be about —348°F. Because it is so very far away, it is extremely difficult to observe. Nevertheless, its period of rotation has been determined by photoelectric measurements by Robert Hardie of the Dominion Observatory, Ottawa. Apparently its surface is not uniform, for the light it reflects varies with a consistent period of 6 days, 9 hours, 17 minutes. This is assumed to be its period of rotation.

The diameter of Pluto remains a bit of a mystery, although its maximum size has been estimated to be 3,600 miles. This estimation is based on the knowledge of which stars Pluto's tiny apparent disk does *not* eclipse as it moves across the sky.

Since Pluto's spectrum does not contain any molecular absorption bands, we assume that it is quite different from the giant planets between Mars and Pluto. In fact, it has been speculated that Pluto was once a satellite of Neptune that escaped and began circling the sun in an orbit of its own. (Since Pluto now revolves about the sun and not about Neptune, it must be classified as a planet and *not*

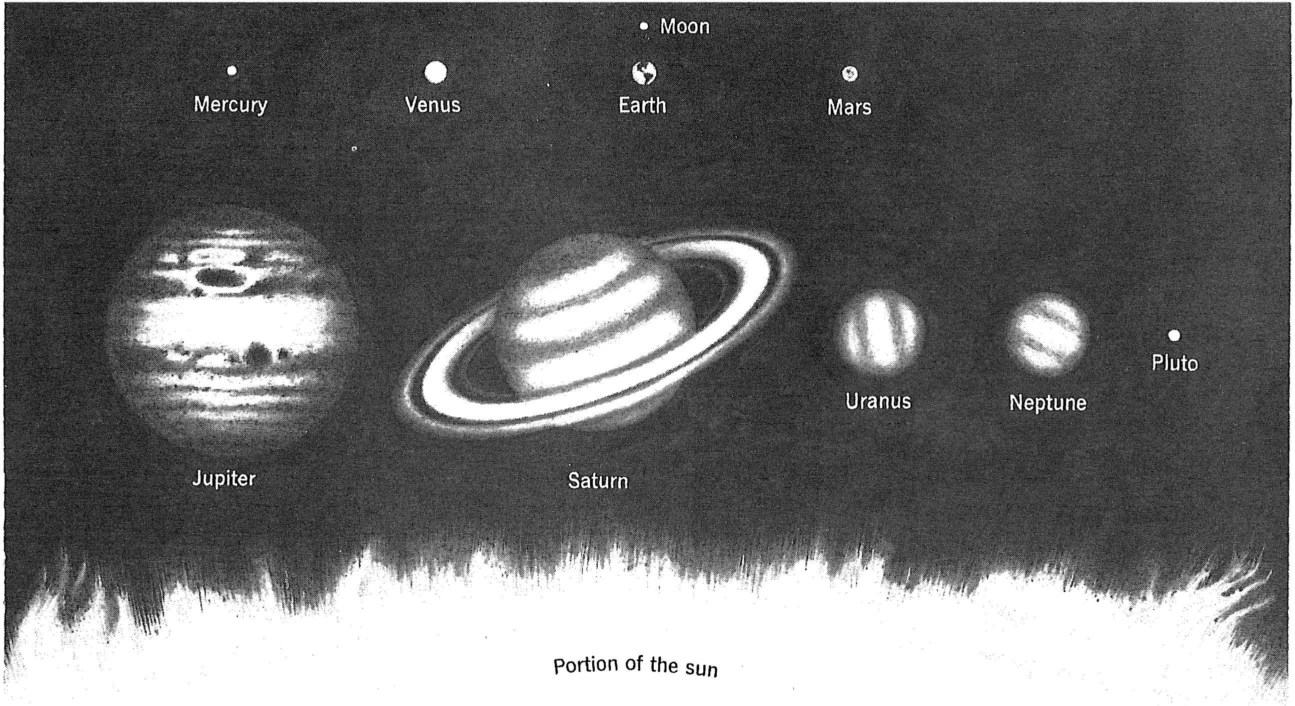


Figure 4-12 The largest members of the solar system drawn to scale. The diameter of the sun is nearly 10 times that of Jupiter, which in turn is nearly 10 times that of the Earth.

as a satellite.) The main justification for this hypothesis is Pluto's eccentric orbit. The planet actually travels inside of Neptune's orbit for a portion of its journey around the sun.

We have spoken of the sizes and distances of the planets, but it is difficult to grasp the dimensions involved because they are so large. Figure 4-12 shows the planets with respect to their sizes. The sun is added in the figure to show how much larger it is than any of the planets (the sun's diameter is nearly 10 times that of Jupiter). This is fairly easy to imagine, but conceiving of their distances is more difficult. Mercury is 36 million miles from the sun and Pluto is about 100 times as far away or 3,662 million miles. This can be drawn to scale, but to fit such a model of the solar system on a page the size of this book it has been necessary to break the solar system into two parts and to draw each to a different scale. This is shown in Figure 4-13. On both scales each planet is too small to be seen except by a good microscope.

The Earth seems large to us; and so, upon reflection, do the other planets. All are complex in their structure and fascinating in their details. But what we know of them, even of those we have studied the most, is little compared to the complete picture we would like to have. It is space and the distance between the planets that keeps us from this knowledge—distances so vast that even Jupiter, 87,000 miles across, is microscopic in comparison to the 480 million miles that separate it from the sun, or the 3,700 million miles that separate the sun from Pluto.

Table 4-1 Data on the Planets

	Mean Distance from Sun							
	Astronomical Units	Million Miles	Period of Revolution	Period of Rotation	Diameter in Miles	Mass in Terms of Earth's Mass	Density in Terms of Water	Escape Velocity, Miles per Second
Mercury	0.39	36	87.96 d.	58.4 d.	3,100	0.054	5.3	2.6
Venus	0.72	67	224.69	−248 d.	7,600	0.82	5.0	6.5
Earth	1.00	93	365.24	23 hr. 56 min.	7,917	1.00	5.5	7.0
Mars	1.52	142	686.95	24 hr. 37 min.	4,200	0.11	4.0	3.2
Jupiter	5.20	484	11.86 yr.	9 hr. 50 min.	88,700	317.8	1.3	37.1
Saturn	9.56	889	29.56	10 hr. 14 min.	75,100	95.2	0.7	22.3
Uranus	19.2	1,782	83.95	10 hr. 49 min.	29,000	14.5	1.6	13.9
Neptune	30.1	2,784	163.9	15 hr. 40 min.	28,000	17.2	2.3	15.4
Pluto	39.4	3,662	247.3	6 d. 9 hr. 17 min.	<3,600	?	?	?

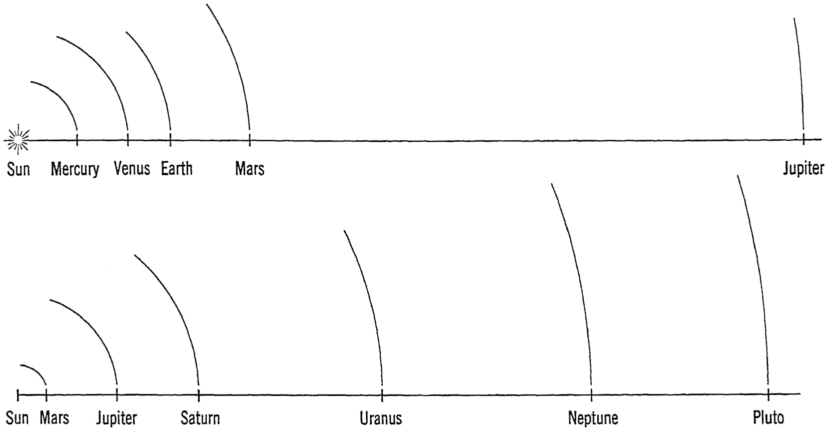


Figure 4-13 Two scale drawings showing the distances involved in the solar system. The upper drawing shows a portion of the orbits of the terrestrial planets and Jupiter, with the asteroid gap between. The lower drawing shows a portion of the orbits of the Jovian planets with Mars for comparison. If any of the planets were drawn on either of these scales, they would be microscopic.

Basic Vocabulary for Subsequent Reading

Albedo
Astronomical unit
Density
Escape velocity
Limb
Mare
Mass
Revolution
Rotation
Scattering of light
Solar system
Terminator
Zenith

For Further Reading

Jackson, Joseph H., *Pictorial Guide to the Planets,* Thomas Y. Crowell Co., New York, 1965.

Moore, Patrick, *The Planet Venus,* Crowell, Collier and Macmillan, New York, 1957.

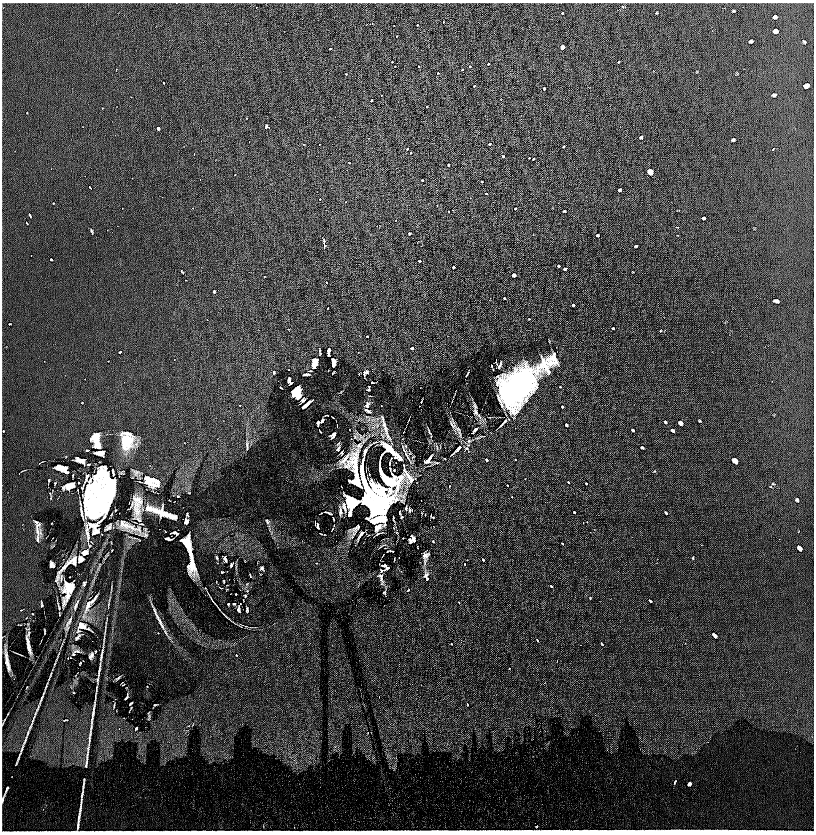
Peek, Bertrand M., *The Planet Jupiter,* Faber and Faber, London, 1958.

Petrie, William, *Keoeeit—The Story of the Aurora Borealis,* Pergamon Press, New York, 1963.

Slipher, Earl C., *A Photographic Study of the Brighter Planets,* Lowell Observatory, Flagstaff, Arizona, and The National Geographical Society, Washington, D.C., 1964.

Smith, Alex G., and Thomas D. Carr, *Radio Exploration of the Planetary System,* D. Van Nostrand Momentum Book, Princeton, N.J., 1964.

Struve, Otto, and V. Zebergs, *Astronomy of the 20th Century,* Chapter VIII, Crowell, Collier and Macmillan, New York, 1962.
Akasofu, Syun-Ichi, "The Aurora," *Scientific American,* p. 55, December 1965.
Barrett, Alan H., and E. Lilley, "Mariner 2 Microwave Observations of Venus," *Sky and Telescope,* p. 192, April 1963.
Cahill, L. J., "The Magnetosphere," *Scientific American,* p. 58, March 1965.
Franklin, Kenneth L., "Radio Waves from Jupiter," *Scientific American,* p. 35, July 1964.
Goodman, Joel W., "The Edgewise Presentation of Saturn's Rings," *Sky and Telescope,* p. 128, September 1965.
James, J. N., "The Voyage of Mariner II," *Scientific American,* p. 70, July 1963.
James, J. N., "The Voyage of Mariner IV," *Scientific American,* p. 42, March 1966.
Leighton, Robert B., "The Photographs from Mariner IV," *Scientific American,* p. 54, April 1966.
O'Brian, Brian J., "Radiation Belts," *Scientific American,* p. 84, May 1963.
Parker, E. N., "The Solar Wind," *Scientific American,* p. 66, April 1964.
Sloan, Richard K., "The Scientific Experiments of Mariner IV," *Scientific American,* p. 62, May 1966.
Strong, John, "Infrared Astronomy by Balloon," *Scientific American,* p. 28, January 1965.
"A Jovian Atmospheric Feature of Special Interest," *Sky and Telescope,* p. 118, February 1965.
"New Radar Studies of Venus," *Sky and Telescope,* p. 356, June 1965.
"New Arecibo Findings About the Planets," *Sky and Telescope,* p. 339, June 1965.
"Pluto's Diameter," *Sky and Telescope,* p. 213, October 1965.



Planetarium projectors can show the complex motions of the planets in a matter of minutes on a domed ceiling with amazingly realistic effects. (Morrison Planetarium, California Academy of Sciences, San Francisco)

Chapter 5 *The Planetary Motions*

Chapter 4 gave us some idea of the physical characteristics of the planets. But this is not a complete description of the solar system, or even of the planets. To complete their story it is necessary not only to discuss their motions but also to mention the historical development of the main ideas that have been proposed to explain those motions.

"Planet"—derived from a Greek word meaning "wandering"—is an appropriate name for these objects in the sky that appear to be continually moving stars. Their motions were observed many thousands of years ago and can be observed today by looking at the sky with some care.

Few people, however, pay much attention to the sky. If they did they would realize that the planets move against the background of stars over the weeks and months. They would also be able to see that the moon moves against the same backdrop in just an hour's time, and that even the sun moves differently in the sky than do the distant stars.

It should be noticed that each of these motions has been related to the background of stars, for these apparent motions become more obvious when compared to the "fixed" system of stars. But that to which we must ultimately refer all these motions is the Earth itself, for the sun, the moon, and the planets would appear to move as they do if the stars were not there. In other words, the Earth (the local horizon) is the *frame of reference* against which we measure all these apparent motions.

Every motion must be referred to some frame of reference. The motion of a car is tacitly referred to the Earth when we say that it is traveling 60 miles per hour, for it is traveling much faster than this when referred to celestial frames of reference.

The apparent *diurnal* (daily) motion of the stars is from east to west. The moon moves differently than the stars in that it moves westward more slowly than they do. Consequently it appears to move

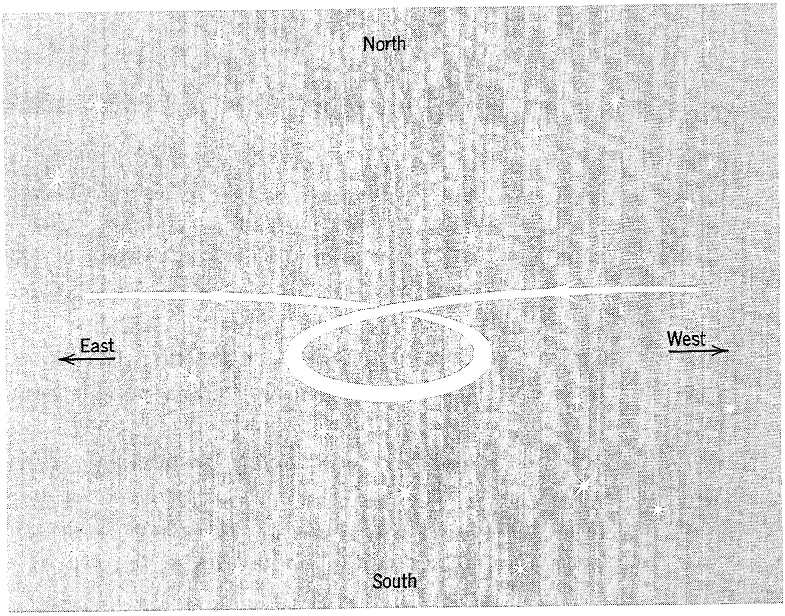


Figure 5-1 The westward motion of Mars (or any planet) is called retrograde motion.

eastward relative to the stars (that is, it lags behind in the "diurnal race") approximately one of its diameters (½°) every hour. Thus the moon sets on the average 49 minutes (1° = 4 minutes of time) later each day.

The sun, too, moves eastward relative to the stars, but this motion is only about 1° each day. To put it another way (since we reckon civil time by the sun), the stars move west relative to the sun about 1° each day. Thus each star sets about 4 minutes earlier on succeeding nights.

The planets, generally speaking, also move eastward through the field of stars, but there are times when each planet moves westward. Since this motion is contrary to the usual motion it is called *retrograde motion.* Figure 5-1 shows schematically how Mars may move during retrograde motion when it is brighter than at any other time.

Kepler's Laws

Since the time of Plato, in ancient Greece, there have been a number of theories to explain the motion of the planets. Plato, in effect, challenged the mathematicians to explain the apparently erratic motion of the planets, and he went so far as to specify that the motion must be circular, at a steady speed, and *geocentric* (from the Greek, *ge,* Earth). The circular orbits and the steady speed were selected essentially because of their basic symmetry and "perfection." The geocentrism, however, was undoubtedly favored because it is more comforting to believe that the Earth is *not* moving, and more pleasing to man's ego to think that the Earth is at the very center of the universe. However, in the seventeenth century, the *heliocentric* (from the Greek, *helio,* sun) theory became established once and for all, but not without a struggle and not without men saying, "Prove it to me!" But theory must precede the proof.

In the first two decades of the seventeenth century, Johannes Kepler published what have become known as Kepler's three laws of planetary motion. The first two laws describe the shape of a planetary orbit and the motion of the planet in that orbit. The third law compares the motion of any two of the planets revolving about the sun. Although the laws were derived specifically for the planets revolving about the sun, they apply to satellites revolving about the Earth and planets revolving about any star.

The shape of each planetary orbit is described by the first law as an *ellipse.* An ellipse is a figure that can be drawn by taking a piece of

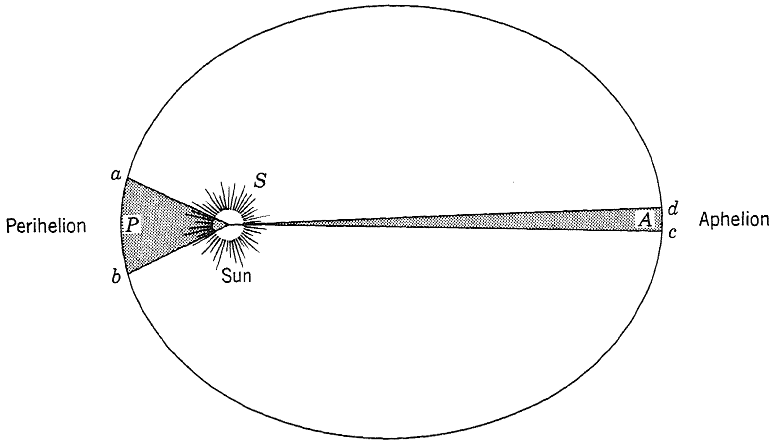


Figure 5-2 According to Kepler's second law, the line joining the sun with any planet sweeps out equal areas in equal intervals of time. If the planet takes the same length of time to move from *a* to *b* as it does from *c* to *d* then the areas *Sab* and *Scd* must be equal.

string, 8 in. long, for example, tacking the ends down by tacks that are placed 6 in. apart, and running a pencil inside the string, always keeping the string taut. The position that each tack occupies is called a *focus. The sun occupies one of the foci of each planetary orbit.*

Since the sun is at one of the foci, the distance of each planet from the sun varies (Figure 5-2). When the planet is in position A, *aphelion,* it is at its farthest from the sun. When it is in position P, *perihelion,* the planet makes its closest approach to the sun. The corresponding terms in the orbit of a satellite of the Earth are *apogee* and *perigee.*

The second law is a little more difficult, but it is important for understanding the motions of objects about a central body when under gravitational attraction. The line joining the planet with the sun is called the *radius vector* and, according to Kepler's second law, *the radius sweeps out equal areas in equal lengths of time.* This means that while the planet goes from point *a* to point *b* (both near perihelion) in a certain length of time, say one month, its radius vector will sweep out the area that is shaded. Starting from point *c,* its radius vector will sweep out the same area in the same length of time, but the planet will move more slowly since the radius vector is longer near aphelion. Thus when a planet is near perihelion it travels faster than when it is near aphelion. The average velocity of the Earth in its orbit is 18.6 miles per second; at perihelion it travels 18.9 miles per second and at aphelion it travels 18.3 miles per second.

Kepler's third law describes how the periods of the planets depend on the average radii of their orbits. It can be most easily expressed in the form of a proportion:

$$\frac{R^3}{r^3} = \frac{P^2}{p^2}$$

where R is the average orbital radius of one planet, r the average orbital radius of the other, and P and p their respective periods of revolution about the sun. In effect, Kepler's third law tells us that the closer a planet is to the sun the faster it travels. Mercury's average orbital velocity is 30 miles per second, that of Pluto is 3 miles per second.

Kepler knew the periods of all the planets visible to the naked eye and he had only to find the length of the orbital radius of one planet to calculate the orbital radius of each planet and thus determine the size of the solar system.

The average distance of the Earth from the sun is 1 A.U., so Kepler's third law permits us to find the distance of any planet in astronomical units by comparing that planet with the Earth. Jupiter's period of revolution is nearly 12 years, the Earth's is 1 year; the distance of the earth from the sun is 1 A.U. We can now find R, the average distance of Jupiter from the sun:

$$\frac{R^3}{1} = \frac{12^2}{1}$$

$$R^3 = 144$$

$$R = 5.2 \text{ A.U.}$$

Although Kepler never learned the size of the solar system in comparison to a standard distance on the surface of the Earth (such as a mile), he did know that Jupiter was about 5.2 times as far from the sun as the Earth is.

During the intervening centuries, the length of the astronomical unit in miles has been determined with increasing accuracy. The most accurate method is to send radar signals with a radio telescope to Venus and record the length of time it takes for their echo to return to that same telescope. The velocity of light is known with a high degree of accuracy, and so the distance of Venus at the moment the radar signals bounce off its surface depends upon our ability to time the interval it takes for the signals to make the round trip to Venus and back. The results of this technique indicate that the length of the astronomical unit is approximately 92,956,000 miles.

Newton's Laws

Newton sought the answer to the questions "Why does the moon travel about the Earth in an elliptical orbit? Why doesn't it fly off into space and travel in a straight line at a constant speed?"

Newton recognized that any object which has no unbalanced force acting on it will either remain at rest or, if it is in motion, will continue its motion in a straight line at a constant speed. This is Newton's *first law of motion.* We sometimes call it the law of inertia, for inertia is that property of matter which resists change from either its position of rest or its motion in a straight line at a constant speed.

Any change in motion we call an *acceleration* whether it is an increase or decrease in speed, or a change in the direction of travel. Newton realized that for an object to accelerate, an unbalanced force has to act on that object. It is Newton's *second law of motion* which explains that the cause of acceleration is an unbalanced force. The larger the unbalanced force acting on an object, the greater the acceleration. But not all objects have the same inertia, that is, the same resistance to acceleration. The amount of inertia that an object has is called its *mass,* and the greater the mass (for the same force) the less the acceleration. This can be expressed mathematically as

$$\text{Force} = \text{mass} \times \text{acceleration}$$

$$F = ma$$

Therefore, since the moon (or an artificial satellite or a planet) is continually accelerating by continually changing its direction of travel, there must be an unbalanced force acting on it (Figure 5-3). Since the moon travels in nearly a circular orbit this force must act so as to pull the moon toward the center of that orbit.

But the Earth is at the center of the moon's orbit, and this fact led Newton to the conclusion that it is the Earth which exerts the required force on the moon. Likewise, he felt, the moon must exert a force on the Earth. But if the moon and the Earth exert an attractive force on each other, then so must the sun and the Earth, the sun and Saturn, the Earth and Saturn, etc. Newton then derived the ex-

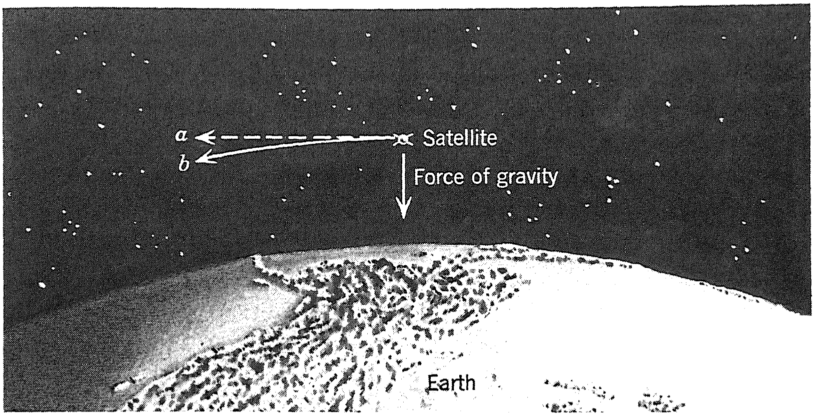


Figure 5-3 Any object, such as a satellite, if not acted on by a force, would continue traveling in a straight line. The Earth's force of gravity. however, causes a satellite to follow an elliptical path. The satellite is always falling but is moving ahead with such a high velocity that it falls into an orbit which misses the Earth.

pression relating the force of gravitational attraction between two bodies to the mass of each body and to the distances separating them. This relationship is *Newton's law of gravity: every object in the universe attracts every other object in the universe with a force that is proportional to the product of the masses of the two objects,* m_1 *and* m_2, *and inversely proportional to the square of the distance* d *between them.*

$$F = G\frac{m_1 m_2}{d^2}$$

where G is the universal gravitational constant which changes the proportionality expressed above into an equality. In effect G tells us something about the universe, something about the strength of the gravitational forces when compared with other forces such as electrical forces.

But how does this gravitational force operate? Newton was not able to say, but the fact that it does operate as described by his law of gravity and that it alone accounted for the motion of the planets (until the introduction of the principle of general relativity some 300 years after Newton) makes it a bold and decisive step in man's description of the universe.

But even if Newton's laws of motion and his law of gravity were generally accepted within a few decades after he published them in 1687, men still looked for observational *proof* that the Earth rotates on its axis and revolves about the sun.

Proofs of Rotation

Let us consider rotation first. The most substantial evidence that the Earth rotates was not brought forth until the nineteenth century, when a Frenchman named Foucault presented an argument based upon Newton's second law of motion: before a body can accelerate, a force must act on the body in the direction of acceleration.

If a pendulum (Figure 5-4) is so constructed that it can swing equally well in all directions, and is then set swinging in one plane, it will continue to swing in that same plane if there is no force to cause it to deviate from it. If such a pendulum were placed on the north or south pole, it would oscillate in a plane perpendicular to the plane of the Earth's equator. If the pendulum had been set in motion in the direction of a certain star, then its plane of oscillation would continue to point to that star as the star appeared to circle the sky around the north pole. (As seen from the north pole the stars do not rise and set but travel in circles parallel to the horizon.) As a result it would seem to an Earthbound observer that the direction of the pendulum's plane of oscillation was changing. If the Earth did not rotate, however, the plane of oscillation of the pendulum would remain fixed with respect to both the Earth and the stars.

By means of this argument and with an actual pendulum, Foucault demonstrated that the apparent changes in direction of swing could be explained only if the Earth rotated on its axis with respect to the stars.

A second proof for the Earth's rotation has been discussed already. This is the oblateness of the Earth. If a body that is able to change its shape over a period of time is acted on by a force that is directed toward the center and is equal in all directions, the body will assume a spherical shape. A balloon is such a body; a small drop of oil under the surface of water is another example.

If the Earth's gravitational force, which is uniform in all directions, were permitted to shape a nonrotating Earth, the Earth would be spherical. But its polar diameter is less than its equatorial diameter. What, then, keeps it from being a perfect sphere?

The only way to explain this oblateness of the Earth is to apply Newton's laws of motion to a rotating Earth. As Newton stated in his first law of motion, all matter that is not acted on by a force

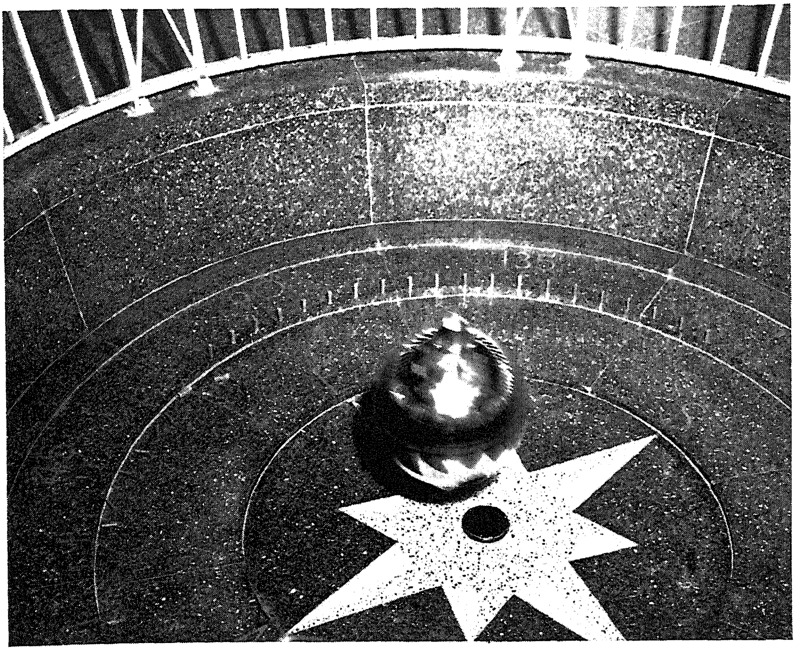


Figure 5-4 The Foucault pendulum at the California Academy of Sciences. The plane of oscillation is rotating as can be seen by the pegs that have already been knocked down by the pendulum. (California Academy Sciences)

will remain at rest or move with a constant velocity in a *straight line.* To make material that is moving with a constant velocity deviate from a straight line, a force must be applied. This is the situation of material on the equator of a rotating Earth. There must be some force that causes this material, which is moving in a circular path, to deviate from straight-line motion. The only force available for this is the force of gravity. Some of the force of gravity is used to make the material on the equator move in a circle, and thus slightly less gravitational force is available to shape the Earth at the equator. The material near the poles, however, is rotating very slowly, so in these regions practically all the gravitational force is available to shape the Earth. Consequently, the Earth's polar diameter is less than its equatorial diameter. No explanation other than that outlined above will account satisfactorily for our planet's oblateness.

Consequences of Rotation

The principal consequence of the Earth's rotation is our day and night. It is a consequence only, and not a proof, because it can be accounted for by assuming that the sun revolves around the Earth. It should be made clear that although every proof is a consequence, every consequence is *not* a proof.

Related to the rotation of the Earth, and indirectly a consequence of it, is the fact that the two points on the Earth where the axis of rotation intersects the surface, called the poles, are not always the same points. That is, the poles wander slightly on the surface of the Earth. This wandering of the poles is not a motion of the axis but of the Earth on its axis. The causes of this wobbling are not quite clear, for the motion is complex and may be the result of many causes all added together. It does seem reasonable to assume, however, that a shifting of mass, such as water in the Earth's atmosphere or perhaps material within the Earth itself, could contribute to the wandering of the poles.

Even the Earth's period of rotation itself is not constant. The day is about $\frac{1}{1000}$ of a second shorter during our northern autumn than in spring. This may be caused by a shifting of water vapor in the Earth's atmosphere. In addition there are other variations of about $\frac{1}{1000}$ of a second in the length of the day that occur over shorter intervals of time. These may be attributed to movements in the Earth's interior.

Proofs of Revolution

It was realized rather early that if the Earth revolved around the sun and if the stars were at different distances, then as a consequence of changes in the Earth's position there should be a shifting back and forth of the nearest stars with respect to those more distant. This apparent annual shift of position for the nearest stars is called *heliocentric parallax* (Figure 5-5).

When the Earth is in the position shown in *a* the nearer star appears in the direction indicated, and a photograph of the region might appear something like star plate 1. In *b*, 6 months later, the Earth has changed its position and the nearer star is seen from a different direction. The corresponding photograph might appear as star plate 2.

The result is that the nearest stars will appear to oscillate back and forth against the background stars in a period that is equal to the period of the Earth's revolution. This fact, since it cannot be explained independently of the Earth's revolving, constitutes in itself sufficient proof that the Earth does revolve.

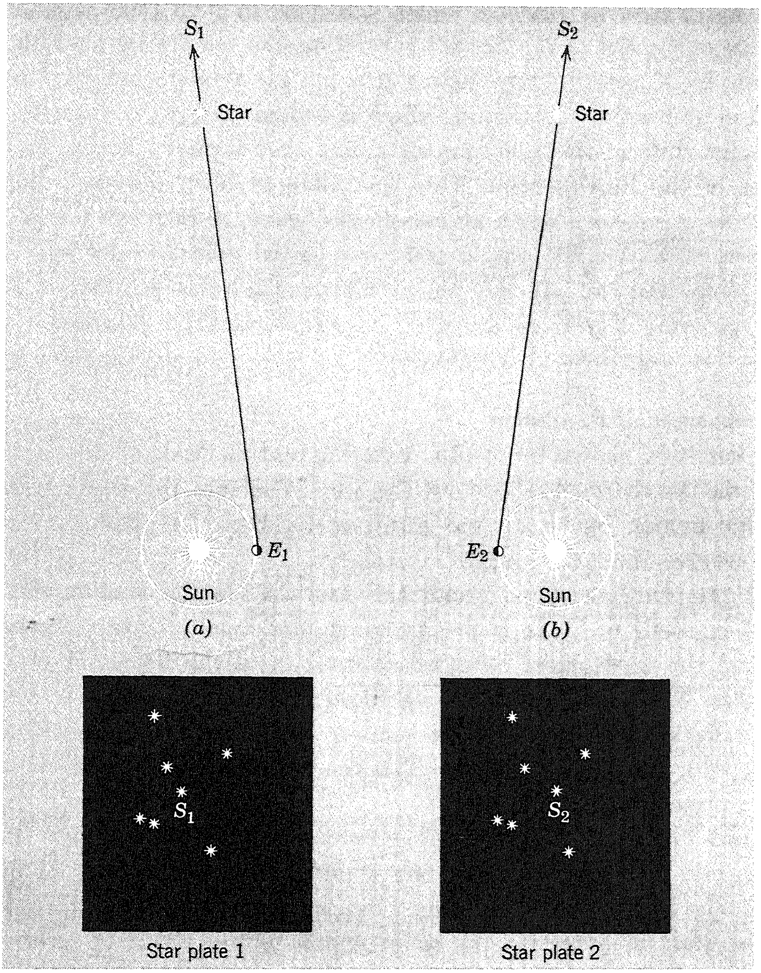


Figure 5-5 At one time of the year a nearby star (*a*) will be seen in one direction, S_1; 6 months later (*b*) the same star will be seen in a slightly different direction, S_2. The difference in direction is highly exaggerated, for in reality it is always less than 1 second of arc. Star plate 1 shows the position S_1 against the more distant stars. Star plate 2 shows the new position of the nearby star.

Since the stars are so far away in comparison with the radius of the Earth's orbit, the resulting change in their positions is extremely small and thus was not detected until the nineteenth century. Other observations before the nineteenth century, however, helped convince many of the philosophers and scientists that the Earth does revolve about the sun. One such observation was the measurement of the velocity of light by Roemer, which was done in 1675 (see page 16f).

A second proof of the Earth's revolution can readily be made from the study of spectroscopy. The Earth has an average velocity in its orbit of 18.6 miles per second, which is enough to cause a measurable Doppler shift in the spectrum of a star that is in or nearly in the plane of the Earth's orbit. This has not only been observed but it becomes a nuisance when an astronomer wants to measure the radial velocity of a star. To obtain consistent radial velocities for any star it is convenient to use the sun as a frame of reference rather than the revolving Earth, so the Earth's motion must be subtracted out of each measurement of radial velocity.

Consequences of Revolution

It has been shown by at least two different methods of observation that the Earth revolves around the sun. What are the consequences of this motion for us on the Earth or for the other planets which also revolve about the sun?

As the Earth revolves about the sun we see the sun appear to revolve about us. This is the apparent motion of the sun eastward against the background of stars referred to in the first part of this chapter. The sun's apparent path through the field of stars is called the *ecliptic*, which is the intersection of the plane of the Earth's orbit with the celestial sphere. The sun, like all celestial objects, appears projected on this sphere.

Since the Earth moves in an elliptical orbit it must rotate more than 360° in order to make one complete 24-hour period (Figure 5-6). The Earth in position *a* must rotate more than 360° with respect to the stars in order for the same region of the Earth to return to midnight in position *b*. Since the Earth revolves about 1° each day, it must rotate 361° with respect to the stars in a 24-hour period. The rotation with respect to the stars yields the *sidereal day;* the rotation with respect to the sun gives us the *solar day*. The length of the sidereal day is about 23 hours, 56 minutes solar time.

Another consequence of revolution is that puzzling retrograde motion of the planets. Since Mercury travels faster than Venus, when they are on the same side of the sun Mercury will overtake Venus. Similarly Venus will overtake the Earth and the Earth will overtake Mars, etc. When we overtake Mars it appears to us as if Mars is going backward, just as passing a car on the highway makes the slower car appear to go backward.

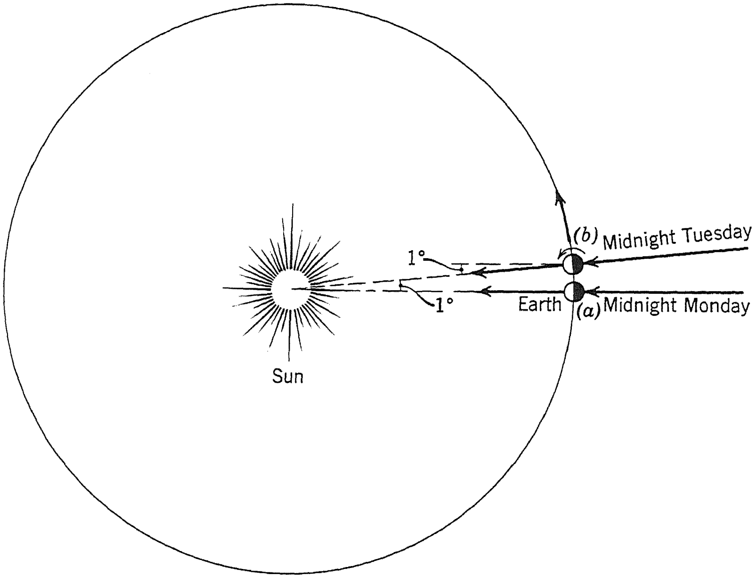


Figure 5-6 During any 24-hour period the Earth moves along its elliptical orbit so that in order for midnight to return to the same locality, the Earth must rotate about 1° more than one complete rotation, or about 361°.

A major consequence of the Earth's revolution is the seasons that we experience: winter, spring, summer, and autumn. It was pointed out that the axis of rotation of the Earth is not perpendicular to the plane of the Earth's orbit (see page 87). For purposes of this discussion only, the axis of rotation may be said to maintain the same direction with respect to the stars (Figure 5-7). When the Earth is in position 1 the north pole is tipped toward the sun, and the south pole is tipped away from the sun. The darkened portion of the Earth is in the shadow and therefore in nighttime. It can be seen that not only is the north pole receiving sunlight, but, as the Earth rotates on its axis, the sun will not set. This is called the midnight sun. At the same time, the south pole is in continual night.

The sun's rays are shining more nearly perpendicularly on the northern hemisphere and so their heating effect is stronger. At the same time they are shining more obliquely on the southern hemisphere and their heating effect is accordingly diminished. Thus when the Earth is in position 1 the northern hemisphere is beginning summer while the southern hemisphere is beginning winter.

Three months later the Earth has moved over to position 2 while

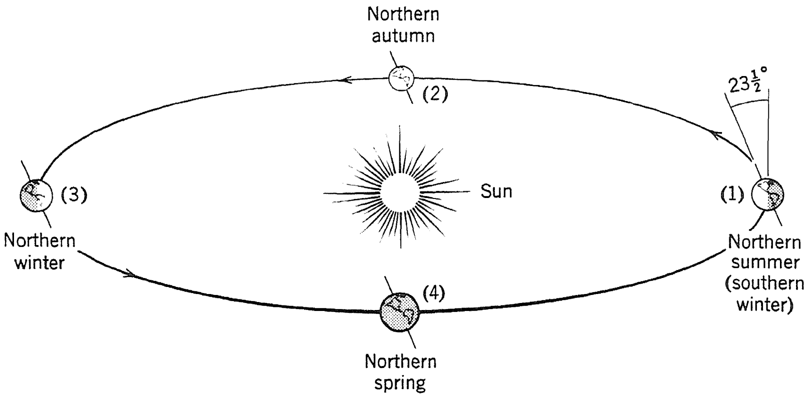


Figure 5-7 The seasons result from the fact that the Earth's axis of rotation is not perpendicular to the plane of the Earth's orbit, and because this axis of rotation points to essentially the same point in the sky during the year.

its axis has remained parallel to the direction in which it pointed in position 1. At this time both hemispheres are equally illuminated by the sun. The sun is on the horizon of both poles, that is, it is setting on the north pole and rising on the south pole. This is the beginning of autumn for the northern hemisphere and spring for the southern hemisphere. Similarly, the seasons follow in sequence as the Earth moves on to positions 3 and 4.

Referring to position 1 we can see that the sun is north of the Earth's equator, and in position 2 it is in the plane of the Earth's equator. In position 3 the sun appears to be south of the equator, and in position 4 it is again directly overhead at noon for those on the equator.

As viewed from the Earth, therefore, the sun appears to move north and south in the sky as well as eastward. The celestial sphere is bisected by the *celestial equator* just as the Earth is bisected by the terrestrial equator; in fact, the projection of the Earth's equator onto the celestial sphere defines the celestial equator. The sun appears to travel throughout the year from the northern celestial hemisphere to the southern celestial hemisphere and back again, crossing the celestial equator twice (Figure 5-8). When the sun crosses the celestial equator on its apparent journey north, about March 21, the northern hemisphere begins spring and the southern hemisphere begins autumn. This point of intersection of the ecliptic with the celestial equator is called the *vernal equinox*. The other intersection of the ecliptic with the celestial equator, that is, where the sun crosses the

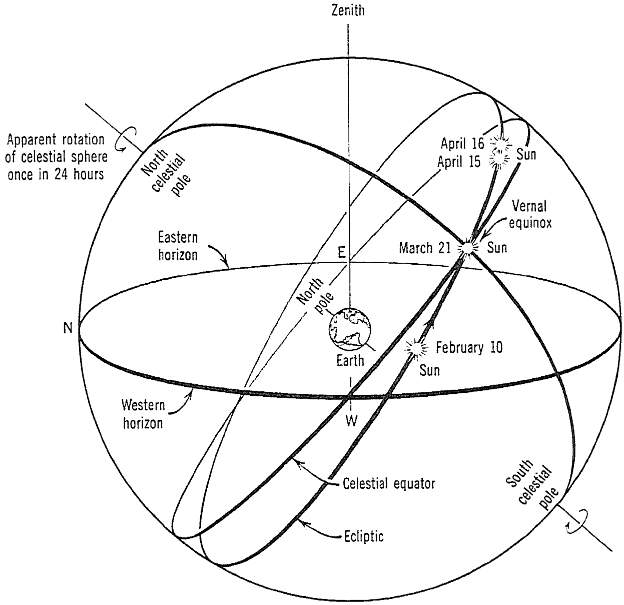


Figure 5-8 The Earth is drawn at the center of a greatly decreased celestial sphere. The Earth's axis of rotation intersects the celestial sphere at the north and south celestial poles. A projection of the plane of the Earth's equator on the celestial sphere is the celestial equator. The apparent path of the sun against the background stars is called the eclipic. The sun moves eastward along the eclipic about 1° each day.

celestial equator going from the northern to the southern hemisphere, is called the *autumnal equinox*. The crossing occurs about September 21.

The point in the sky where the sun reaches its most northern position, about June 21, is called the *summer solstice*. Its most southerly point, the *winter solstice*, is reached about December 21.

It should be pointed out that since these points were named by inhabitants of the northern hemisphere, they naturally refer to the seasons in that hemisphere. The seasons are exactly opposite for people in the southern hemisphere. When the sun is in the vernal equinox, autumn begins for those south of the equator, and so forth.

Great Circles in the Sky

It becomes essential for astronomers to have a coordinate system for defining the position of each star in the sky, just as the position of a given locality on the Earth is defined by its latitude and longitude.

Defining a position on the Earth requires two dimensions: one north or south of some reference circle on the Earth, and one east or west of some other reference circle. The reference circle for the first is quite obvious, for the Earth is divided into two equal hemispheres by the equator, which therefore has become a prime reference circle.

Any circle on the Earth parallel to the equator is called a *parallel of latitude* and defines a given locality as being so many degrees north or south of the equator. The latitude of San Francisco, for example, is about +38° (the plus sign referring to northerly latitudes). The latitude of Sydney, Australia, is about —34° (the minus sign signifying a southerly latitude) and the latitude of the north pole is +90°.

The reference circle for the east and west measurement is not so obvious. Any circle on the Earth that passes through both poles is called a *meridian.* It had to be agreed that one of the meridians should be chosen as the reference and called the *prime meridian.* The one selected by international agreement goes through the Greenwich Observatory in England. Any locality on the Earth can, therefore, be referred to by its longitude as being so many degrees east or west of the prime meridian. San Francisco has a longitude of 122° west; the longitude of Sydney is 151° east. East and west longitudes meet on the 180th meridian which goes north and south through the Pacific Ocean just west of Midway Island and east of New Zealand.*

In the sky, positions of the stars also require two dimensions and are most conveniently stated in terms similar to latitude and longitude. As has been already mentioned, the projection of the plane of the Earth's equator intersects the celestial sphere in a circle called the celestial equator. A star's position north or south of the celestial equator is given as so many degrees of *declination,* this being analogous to latitude. A star with a declination of +38° will pass directly overhead for the people in San Francisco. A star with a declination of —34° will pass through the zenith of Sydney.

The choice of the celestial equator as a reference circle is as obvious as that of the terrestrial equator, but, again, we are faced with the choice of a reference circle for the east-west direction in the sky. Circles on the celestial sphere analogous to meridians on the Earth pass through both celestial poles and are called *hour circles.* Since hour circles are to stars in the sky what meridians are to cities on the Earth, the hour circles must rise and set with the stars. The hour circle chosen as the reference is the one that passes through the vernal equinox. A star's position with reference to this prime hour circle is given as *right ascension* and is measured only to the east in hours instead of in degrees (24 hours equals 360°).

---

* With the advent of manmade satellites, it has become convenient to give the longitudes in only one direction from the prime meridian, namely east. The longitude of Sydney would remain the same on this system, but that of San Francisco would become 238°.

For example, the bright star named Vega has a position of right ascension = 18½ hours, declination = +39°. It is therefore located 39° north of the celestial equator and will cross the meridian of San Francisco 1° north of its zenith. It is also located 18½ hours east of the vernal equinox, that is, it will cross the meridian of a given locality 18½ hours after the vernal equinox crosses the same meridian. Antares has a right ascension of 16½ hours and a declination of —26°. It will cross the meridian of Sydney 16½ hours after the vernal equinox and be 8° north of the zenith of Sydney.

Since the astronomer is concerned with the stars and since the stars have an apparent rotation about the Earth different from that of the sun, he uses clocks whose rate is different from those of the clocks used for ordinary timekeeping.

It has been mentioned that the stars set 4 minutes earlier each day according to solar time (see page 97), and therefore the clocks set to *sidereal time* (star time) must run 4 minutes faster each day. The solar day begins at midnight when a point in the sky exactly opposite the sun crosses the meridian of a given locality—or in actual practice crosses the midpoint of a time zone. The sidereal "day" begins when the vernal equinox crosses the meridian of a particular locality (time zones not being used or being of any use in this case).

All this is further complicated by a third motion of the Earth that prevents the vernal equinox from being fixed in the sky.

Precession

The vernal equinox, being one of the intersections of the celestial equator with the ecliptic, depends on the location of both of these circles with respect to the stars. If we use the ecliptic and the stars as a frame of reference, we find that the two equinoctial points move against the backdrop of stars.

The celestial equator is the circle that is everywhere 90° from the

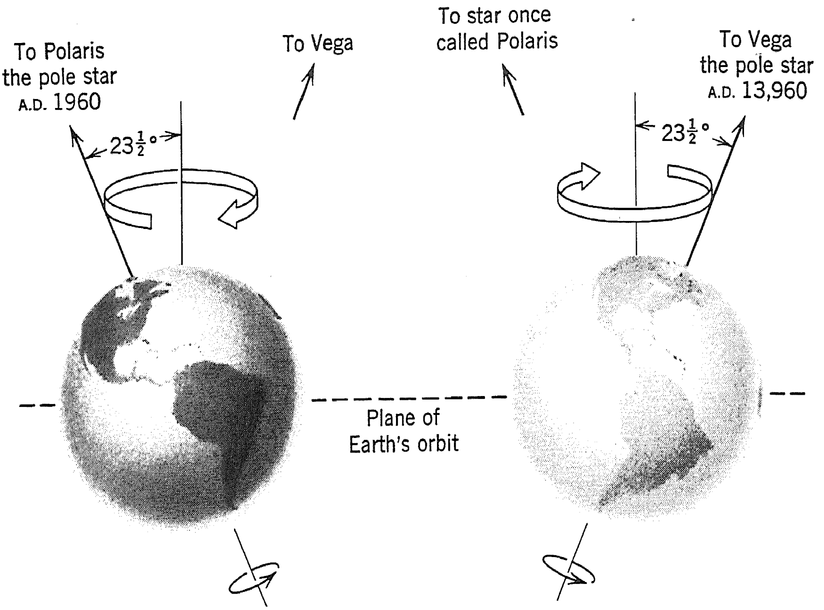


Figure 5-9 The Earth's axis of rotation points nearly to the star we now call Polaris, because of its unique position in the sky. As the Earth's axis of rotation precesses (about a line perpendicular to the plane of the Earth's orbit) the celestial poles will trace out circles on the celestial sphere. In about 12,000 years the north celestial pole will be about 5° from the star Vega.

celestial poles and thus depends on the location of these poles which are the intersections of the axis of rotation of the Earth with the celestial sphere. It is the motion of the Earth's axis itself (not of the Earth about its axis) about a line perpendicular to the plane of the Earth's orbit that causes the celestial equator to move in the sky. This motion, called *precession* (Figure 5-9), is caused by the gravitational pull of the sun and the moon on the equatorial bulge of the Earth. This pull tries to make the plane of the equator more nearly coincident with the ecliptic, but the Earth's rotation will not permit this. As a consequence, the direction in which the axis points continually changes. The axis cuts out a cone in space as it makes one complete rotation in its period of 26,000 years. Since precession would not exist if the Earth did not rotate, the fact that it does precess constitutes proof of the Earth's rotation.

The star we call Polaris is now the north star because the Earth's axis nearly lines up with it, but because of precession it will not always be the north star (see page 6). In 12,000 years the Earth's axis will point to another star, the star Vega referred to above (right ascension = 18½ hours and declination = +39°). It will then be the north star; its declination will be about +85°. Polaris, however, will again be the north star 26,000 years from now. The result of precession, then, is that the positions (right ascension and declination) of all the stars are continually changing. Hence when a position is given the date or *epoch* must also be specified.

These three motions of the Earth are rather difficult to conceive, and may be more easily pictured by likening the Earth to a spinning top placed on the edge of a merry-go-round. The top's spinning represents the rotation of the Earth on its axis, and the top's motion around the center of the merry-go-round represents the revolution of the Earth around the sun. This top, however, is not spinning upright; its axis of rotation is tipped from the perpendicular to the floor of the merry-go-round, with the result that the axis of rotation itself rotates about that perpendicular (as all tops do, especially as their spinning slows down). This motion is precession.

It must be pointed out that all planets experience these three motions to a greater or lesser degree. All three of them may be conveniently referred to the supposedly fixed stars as a frame of reference. Moreover, we have assumed so far that the sun does not travel through space among these "fixed" stars, just as a merry-go-round remains on the same piece of ground.

Basic Vocabulary for Subsequent Reading

Aphelion
Autumnal equinox
Celestial equator
Celestial pole
Celestial sphere
Declination
Diurnal
Ecliptic
Ellipse
Frame of reference
Heliocentric parallax
Latitude
Longitude
Meridian
Perihelion
Precession
Retrograde motion
Right ascension
Summer solstice
Vernal equinox
Winter solstice

For Further Reading

Cohen, I. Bernard, *The Birth of a New Physics,* Doubleday Anchor Books, New York, 1960.

Geymonat, Ludovico, *Galileo Galilei,* McGraw-Hill Book Co., New York, 1965.

Hawkins, Gerald S., *Stonehenge Decoded,* Doubleday and Co., New York, 1965.

Keston, Hermann, *Copernicus and His World,* Roy Publishers, New York, 1945.

Koestler, Arthur, *The Watershed,* Doubleday Anchor Books, New York, 1960.

Kuhn, Thomas S., *The Copernican Revolution,* Modern Library Paper back, New York, 1959.

Munitz, Milton K., *Theories of the Universe,* The Free Press, New York, 1957.

Santillana, Giorgio de, *The Crime of Galileo,* The University of Chicago Phoenix Books, Chicago, Ill., 1959.

Christianson, John, "The Celestial Palace of Tycho Brahe," *Scientific American,* p. 118, February 1961.

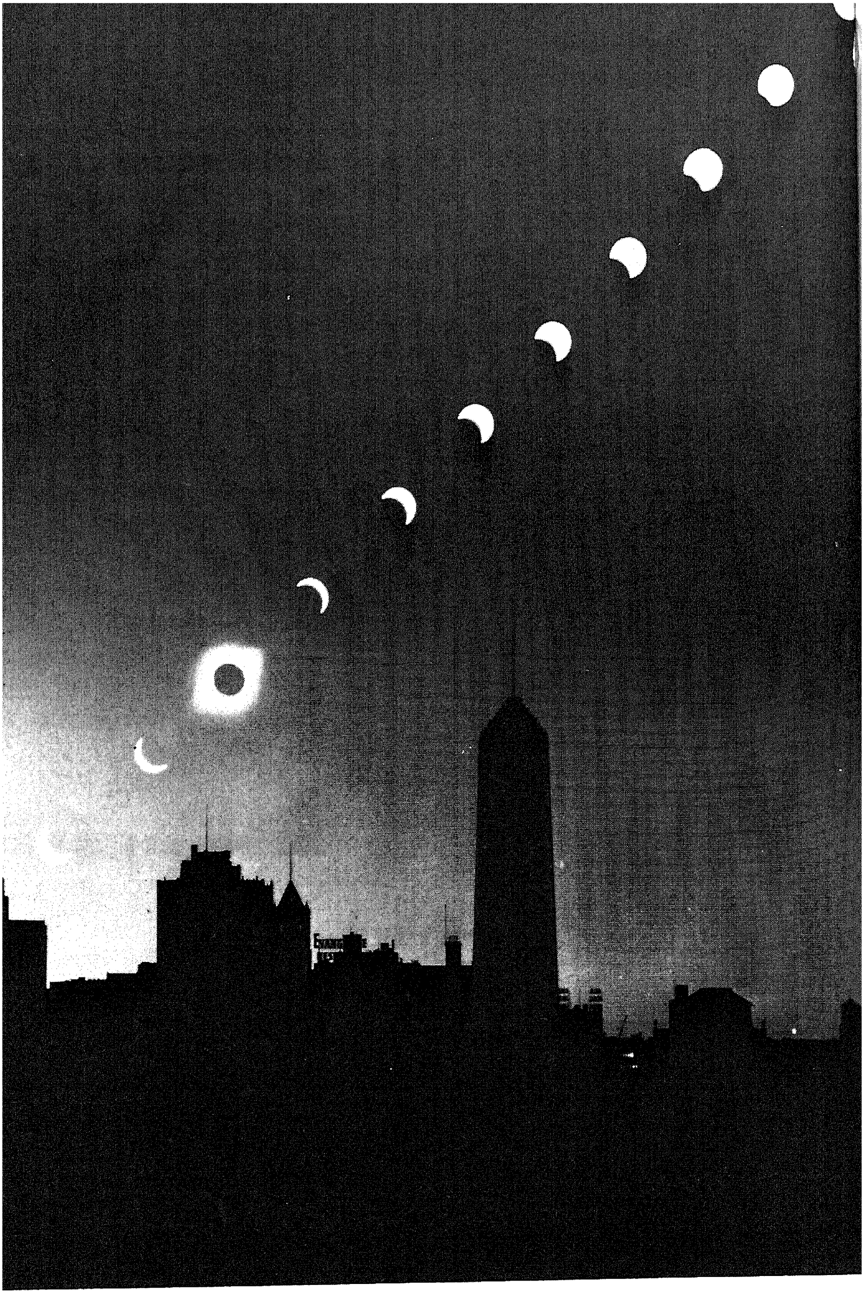
Franklin, Kenneth L., "The Astronomer's Odd Figure 8," *Natural History,* p. 8, October 1962.

Franklin, Kenneth L., "Gravitational Forces and Effects," *Natural History,* part I, p. 12, October 1963; part II, p. 44, November 1963.

Ronan, C. A., "Phoenix of Astronomers," *Natural History,* p. 52, January 1965.

Rothrock, G. A., "Steps to New Astronomy," *Natural History,* p. 64, May 1965.

"Galileo Galilei," series of articles, *Sky and Telescope,* February 1964.



A multiple exposure of the total solar eclipse of June 30, 1954, photographed by Roy Swan. The sun rose in partial eclipse, but as the moon moved eastward relative to the sun it soon caused a total eclipse. The continued motion of the moon progressively uncovered more of the sun, as the rotation of the Earth caused each to rise higher in the sky. (Minneapolis Star Photograph)

Chapter 6 *Satellites*

Our discussion of the planets has covered one phase of the description of the solar system. But of the nine planets, six are known to have satellites circling them as each pursues its complicated motions through space.

Every satellite, since it revolves about a planet, in turn has one more motion than the six characteristic of the planets. For us here on the Earth the most obvious satellite in the universe is our own natural one, the moon, which has an average distance from us of 239,000 miles, and which we see going through its phases during the month.

The Moon

It is not completely correct to say that the moon revolves about the Earth, for in reality they both revolve about a common point

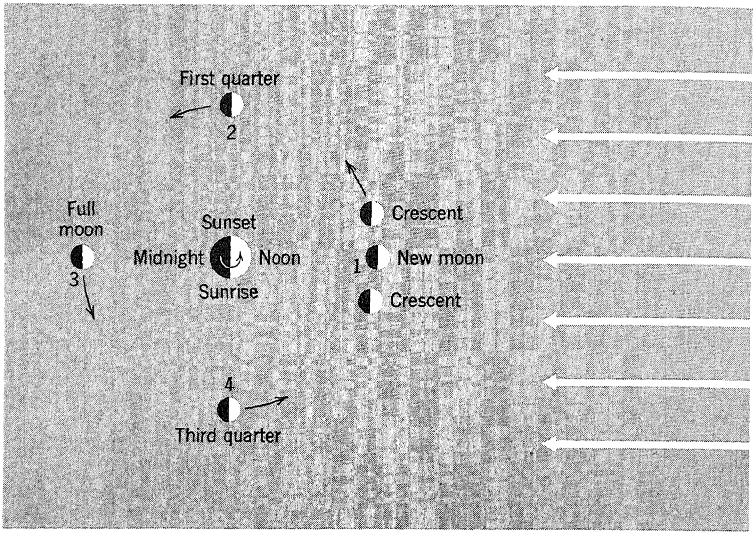


Figure 6-1 **The phases of the moon.**

called their *center of mass.* The Earth is about 81 times as massive as the moon, and the center of mass is therefore about 3,000 miles ($\frac{1}{81}$ of 239,000 miles) from the center of the Earth on a line joining that center with the center of the moon. Since the center of mass is inside the Earth, the Earth's motion about this center is barely perceptible. The corresponding motion of the other planets is even less noticeable, for all other satellites are much smaller when compared with their mother planets.

Since we see the moon only by reflected sunlight, it is understandable that its phases result from its revolution about the Earth (Figure 6-1). Starting with the new moon and following it through the lunar month, the phases progress in order: new moon, first quarter, full moon, and third quarter. Starting with each new month, the length of time within the lunar month is reckoned by the "age" of the moon. That is, the first-quarter moon is said to be 7 days old, etc.

Just before and after the new moon, the sunlit side of the moon visible from the Earth assumes a crescent shape; at first quarter it assumes the half-moon shape. Just before and after full moon, the lighted portion assumes a gibbous shape.

The first-quarter moon differs from the third-quarter in that the western portion of the near side of the moon (as viewed from the

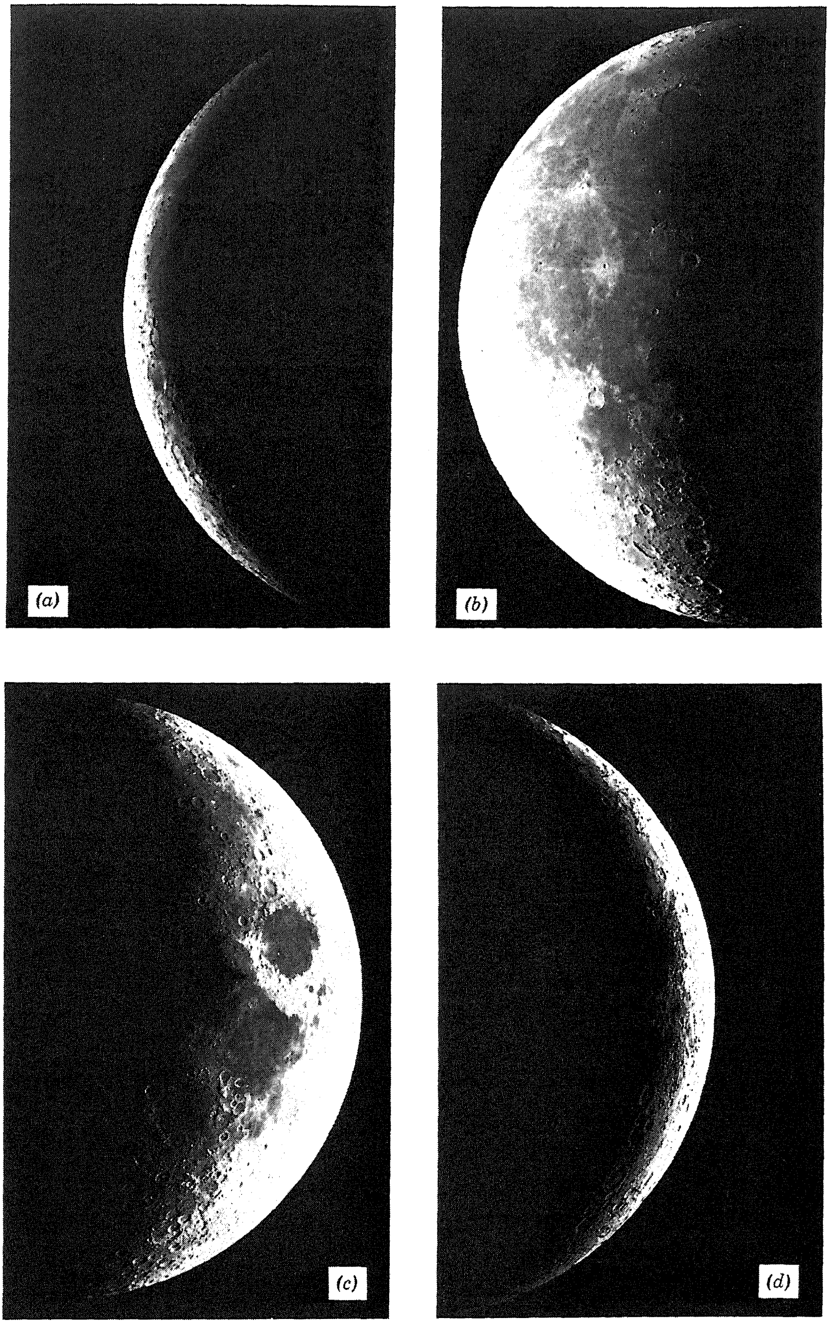


Figure 6-2 (*a*) 26 days old, (*b*) 23 days old, (*c*) 4 days old, (*d*) 3 days old. (Mount Wilson and Palomar Observatories)

Earth) is visible during first quarter; the eastern portion is visible in the third-quarter moon. The first-quarter moon rises about 6 hours after sunrise, so it is high in the sky at sunset. The third-quarter moon rises 6 hours after sunset, so it is high in the sky at sunrise.

The crescent moon has one unique feature. Most of the side that faces the Earth is dark. But the side of the Earth that faces the crescent moon is fully lighted; consequently the moon is faintly illuminated by the Earth's reflected light, or *earthlight.* It is by this reflected light that astronomers have been able to estimate the albedo of the Earth at about 0.4. The fact that this is less than the albedo of Venus (0.76) is in line with what we might expect, since the atmosphere of the Earth has fewer clouds than that of Venus; on the other hand it is more cloud-laden than the atmosphere of Mars, with its albedo of 0.15.

What about the moon with its albedo of 0.07, which has already been mentioned in our discussion of Mercury (see page 64)? The spectroscopic test, when applied to the moon, reveals no more atmosphere than there is on Mercury. But the moon's closeness to the Earth enables us to employ a still more conclusive test.

As has been seen, the moon moves eastward through the field of background stars one of its diameters every hour. In the course of its motion it passes in front of stars, blocking them from our view; this is called an *occultation.* If the moon had any atmosphere it would absorb and scatter some of a star's light. Thus just before a star becomes occulted it would appear slightly fainter than it does when not disturbed by the moon. But with the aid of a sensitive photocell, it has been demonstrated that the star's light does not vary until the solid edge of the moon's limb cuts it from our view. This very sensitive test shows, as do all other tests, that the moon has no permanent atmosphere. There may be small pockets of gas that form from time to time; but these, like those that might form on Mercury, would be dissipated into space, for the moon's escape velocity is only 1.5 miles per second (even less than that of Mercury). The only molecules that could possibly form an atmosphere at all would be those of some of the heavier inert gases such as xenon and krypton, but these elements are very rare.

Having thus established that the moon has very little or no atmosphere, we can take an albedo of 0.07 to indicate lack of an atmosphere.

It is the moon's lack of an atmosphere that enables us to distinguish its features so clearly. Even when seen by the naked eye, it does not appear to have a uniform surface. Its darker portions not only provide children with an image of the "man in the moon," but they also indicate differences in light-reflecting properties of the surface. When the full moon is viewed or photographed through a telescope these differences may be clearly seen. When the first- or third-quarter moon is viewed through a telescope irregularities not only in surface materials but also in depth are visible along the *terminator*, the boundary between day and night.

Lunar Surface Features

As Galileo's telescope opened up a new era for astronomical investigation, so has rocket propulsion and modern technology opened up a new and exciting era. The successful space voyages which have photographed the surface of the moon in detail have—as did Galileo's

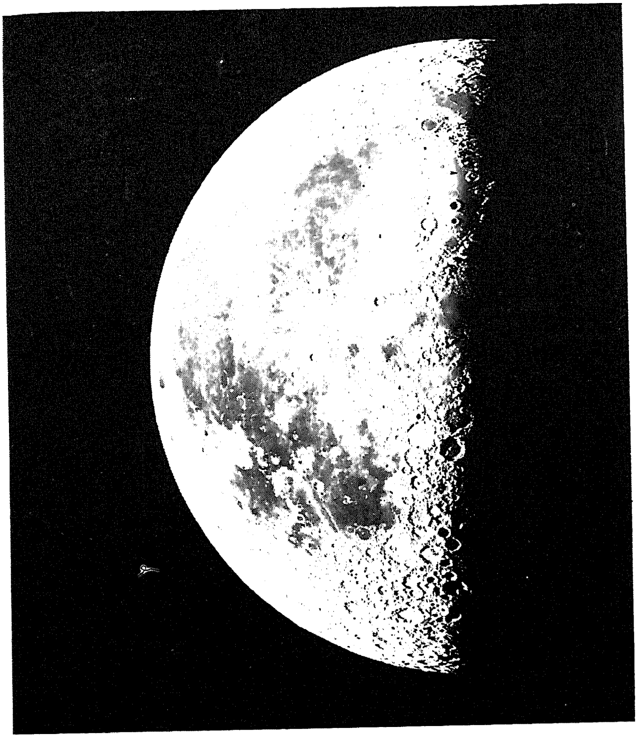


Figure 6-3 The 22-day-old moon. The North pole is at the top of the photograph. (Lick Observatory)

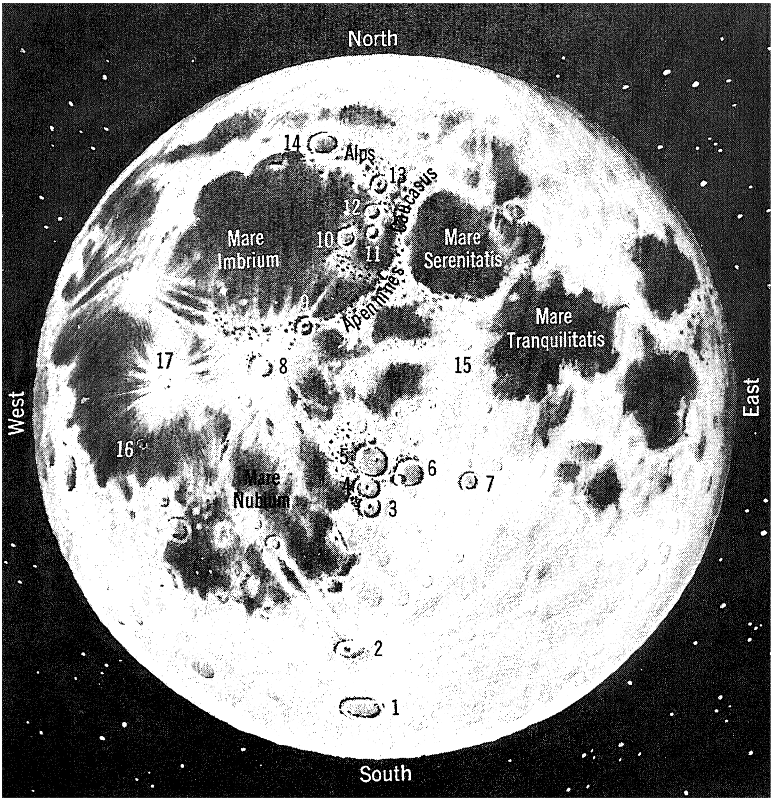


Figure 6-4 A map of the moon. 1. Clavius; 2. Tycho; 3. Arzachel; 4. Alphonsus; 5. Ptolemaeus; 6. Albategnius; 7. Abulfeda; 8. Copernicus; 9. Eratosthenes; 10. Archimedes; 11. Autolycus; 12. Aristillus; 13. Cassini; 14. Plato; 15. Julius Caesar; 16. Flamsteed; 17. Kepler.

telescope—raised more questions than they have answered. But one thing seems to have been settled: the surface of the moon shows evidence of both meteoritic impact and volcanic activity. Furthermore it appears that the lunar topography has been shaped by a number of other causes, some similar to agents which operate to mold the surface of the Earth.

Identification and study of lunar features still proceeds with Earth-based equipment even if space travel has permitted close-up photographs. Much can be done on Earth by the use of radio techniques, large-scale studies of the moon in color, etc. For example, by comparing ultraviolet photographs of the moon with infrared photographs, color differences are made more obvious. This comparison has led astronomers to conclude that the *maria*—the large dark areas—are extensive lava flows. During an extended stage in the moon's past we can only surmise that large amounts of molten rock issued from openings in the crust to fill up the lower portions of the moon. Lava flows have occurred at various times in the moon's history, and, as on the Earth, lunar lava flows differ slightly in their color; some show up better with red filters and red-sensitive film, some better with blue filters and blue-sensitive film.

That lunar history is long and complex is made evident by an inspection of the photograph reproduced in Figure 6-5. This photograph, taken with the 120-in. telescope at the Lick Observatory, reveals portions of old and young lunar topography. The region photographed is the eastern portion of Mare Imbrium. (See Figure 6-4 for identification of some of the lunar features.) North, south, east and west are used here as they would be used by an astronaut on the moon. The highlands or mountainous area to the southeast (lower right) is considered old lunar topography; it predates the formation of Mare Imbrium. Archimedes, the large crater with the smooth floor, must also predate this mare. It would appear that lava flowed out from under the lunar surface through fissures, some even inside crater Archimedes.

The other two large craters in this photograph, Aristillus and Autolycus, were formed after the mare. Evidence of their violent formation may be seen by the whitish *rays* extending radially outward from each crater. The rays even cover the floor of Archimedes, indicating clearly which came first.

There are any number of craters on the moon which have a similar ray pattern. Some of these ray patterns overlap, permitting comparative studies. Photographs from the Ranger series of moon probes demonstrate that in each case of ray overlap, the ray which lies on top of and covers a ray from another crater is brighter than the ray it covers. It would appear that rays become darker with age. This is substantiated by observations of the craters which formed these rays. The craters with the brightest rays have the more complex, rougher-textured walls. Consequently, it must be assumed that the walls of craters, and all portions of the moon, are smoothed by some form of erosion. Ranger photographs reveal that ray systems are accompanied by small craters, called *secondary craters*, caused by mate-

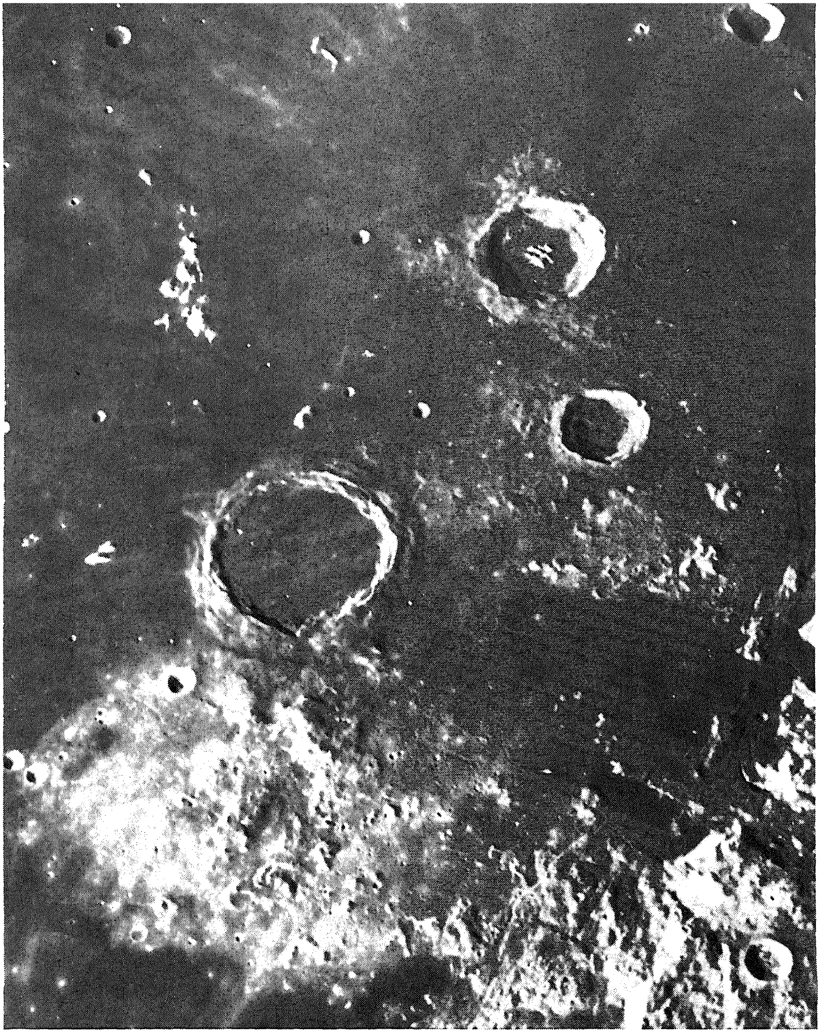


Figure 6-5 The craters Archimedes (the largest one 50 miles in diameter), Aristillus, and Autolycus on the edge of Mare Imbrium. In the lunar photographs, north is up and west is to the right, to correspond to lunar mapping. (Lick Observatory)

rial ejected by the explosion which formed the primary crater. The complex nature of this ray pattern and secondary craters has led astronomers to suspect that craters with rays may have been formed by the collision of a comet with the moon. The rays themselves show evidence of excessive gases which must have originated with the

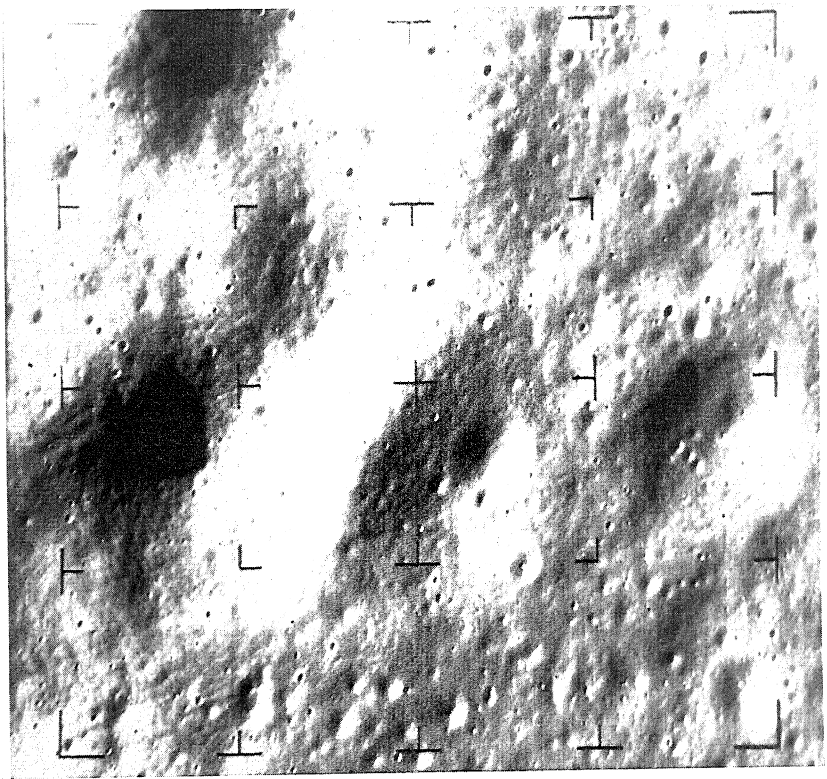


Figure 6-6 Photograph taken from Ranger IX 8.3 miles above the lunar surface and only 5.5 seconds before impact. The surface appears relatively smooth and undulating. (NASA Jet Propulsion Laboratory)

comet and not the moon. These gases literally blew the debris from the primary crater, forming the rays system.

The lava flows that form the maria and floors of many craters appear as smooth surfaces with occasional craters and ridges. This is evident from photographs made with earthbound telescopes (Figure 6-5), as well as from photographs taken by Ranger IX only seconds before it crashed into the surface of the moon in the crater Alphonsus (Figure 6-6). The surface appears smooth, at least in photographs that do not resolve features less than a foot or so in size.

Photographs of the maria made by soft (nondestructive) landings of lunar probes indicate that maria are not composed of sharp jagged rocks, but of smoothed-over surfaces that undulate gently (Figure 6-7). On June 6, 1966, Surveyor I settled gently on the lunar surface

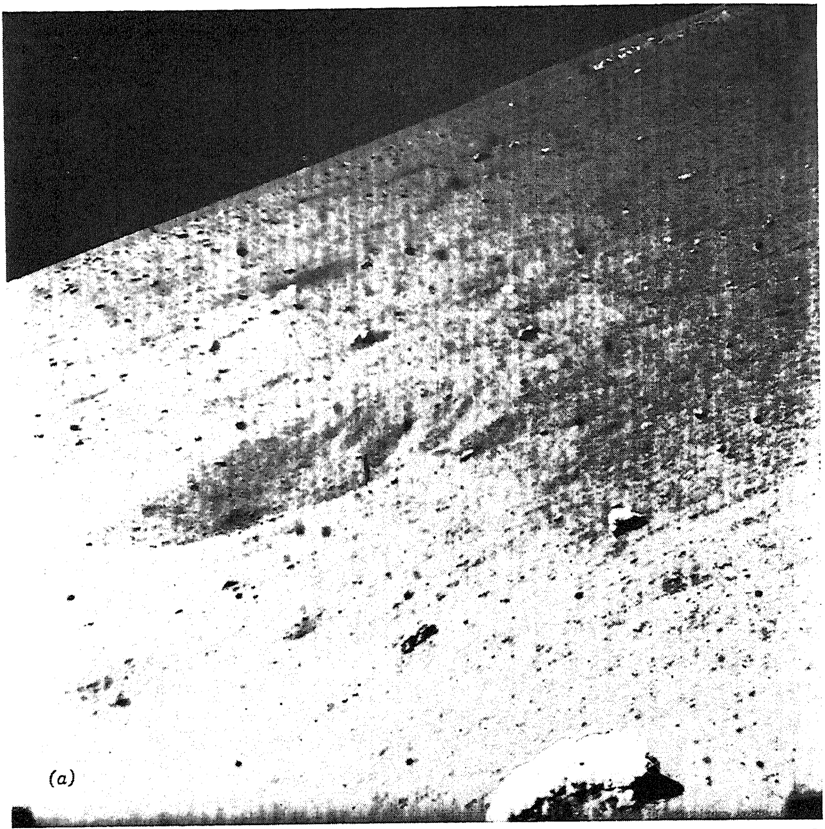


Figure 6-7a The rock strewn surface of the moon photographed with Surveyor I.

on the floor of the crater Flamsteed, a crater that seems to have been filled in with material, since only the tops of its surrounding walls extend above the main plain of Oceanus Procellarum. For the following two weeks Surveyor I took thousands of photographs and radioed them back to Earth. The first impression from these photographs is that the lunar surface is smooth and littered with rocks. There are many small craters, the one to the left of center in Figure 6-7*a* is about 3 yards in diameter. The line of rocks in the upper right-hand corner appears to be a portion of the rim of a much larger crater.

The three landing pads, or feet, of Surveyor I sank into the surface about 1 in., tossing some of the fine grained surface material out

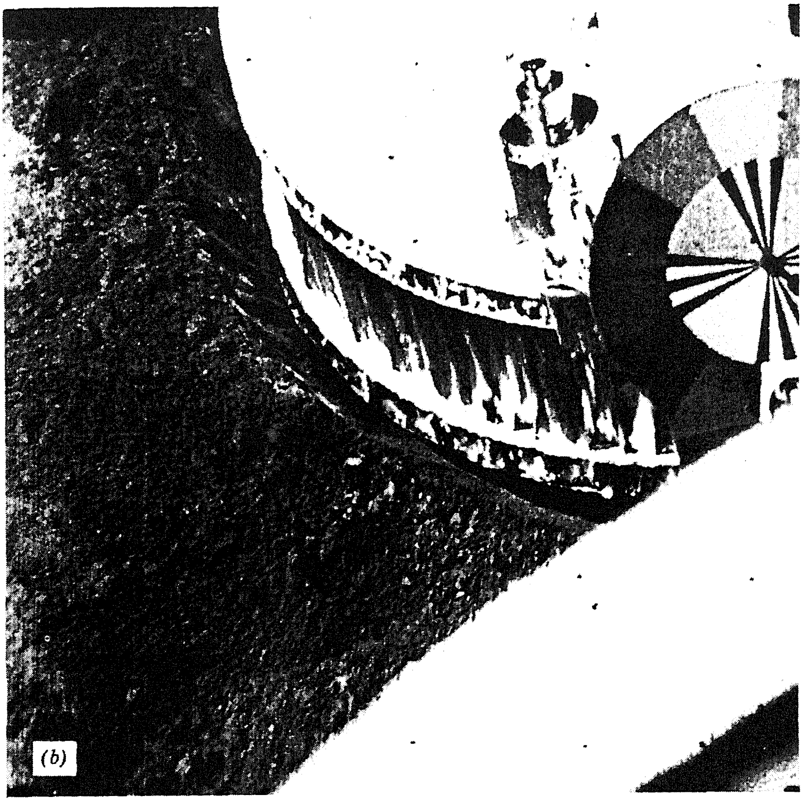


Figure 6-7*b* **The fine-grained surface material of the lunar surface is clearly visible. (NASA Jet Propulsion Laboratory)**

of the hole (Figure 6-7*b*). The material which has been freshly exposed is noticeably darker than the undisturbed surface. Perhaps this change in color is due to the solar radiation. It is supposed that the surface of the moon is eroded by the direct exposure of the surface to the sand-blasting effect of small meteorites and the bombardment of the surface by X-rays and high-energy protons and electrons.

Many of the small craters in the maria appear to result neither from the impact of meteorites nor from volcanic activity. We are led to believe that many craters have occurred because material just under the crust of the lava flow became redistributed to such an extent that portions of the crust, unsupported, collapsed. *Collapse craters* are found in terrestrial lava flows. The crust of a lava flow is the first portion to harden. This permits the lava under the crust to continue to flow and redistribute itself, leaving portions of the crust without support.

That the collapse craters on the moon are associated with *rills,* long troughs on the lunar surface, became clear from photographs of the Ranger series. Observations even with Earthbound telescopes indicate that the floors of some craters are crisscrossed by what appears to be a fairly complex system of rills. Ranger IX hit the surface of the moon near one of these rills (see the white circle in Figure 6-8*a*). As Ranger IX approached the moon, photographs transmitted back to Earth revealed that these rills are the result of many many collapse craters joining each other. Photographs taken by the space craft only 12 miles from the floor of Alphonsus reveal individual craters (Figure 6-8*b*).

Photographs from the Ranger spacecraft reveal many other fascinating aspects of the lunar surface. For example, it is clear that the floor of Alphonsus has a more complex structure and is more crater-riddled than the nearby Mare Nubium (Figure 6-8*a*). If we presume that some of these craters are formed by meteoritic impact, then we can only conclude that the floor of Alphonsus is a good deal older than the surface of Mare Nubium.

It has been suspected from Earthbound observations that there has been recent activity of some sort or another around the crater Alphonsus. Spectrograms taken by the Russian astronomer Kosyrev indicate the occasional presence of luminous gases, for *bright* spectral lines appeared superimposed on the solar spectrum reflected from the lunar surface. It is supposed that the gas is occasionally emitted from the surface and caused to emit its own light by solar radiation. Ranger IX photographs reveal that gases are probably not, however, emitted by the central peak extending 3,500 feet above the surrounding floor, for this peak has no vents through which gases might escape. The peak appears to be very smooth, with little or no detail—at least gross detail. It is possible, of course, that the surface has many small holes through which gases could escape.

That material has been ejected from the surface in the not-too-distant past is made evident by a half-dozen or so craters all on the periphery of the floor of Alphonsus. These unique craters are all surrounded by dark haloes (Figure 6-8*a*). Fortunately, Ranger IX obtained a good photograph of some of these dark-haloed craters. Figure 6-9 is a photograph of the eastern edge of the floor of Alphonsus on the side opposite Mare Nubium. Inspection of the photograph

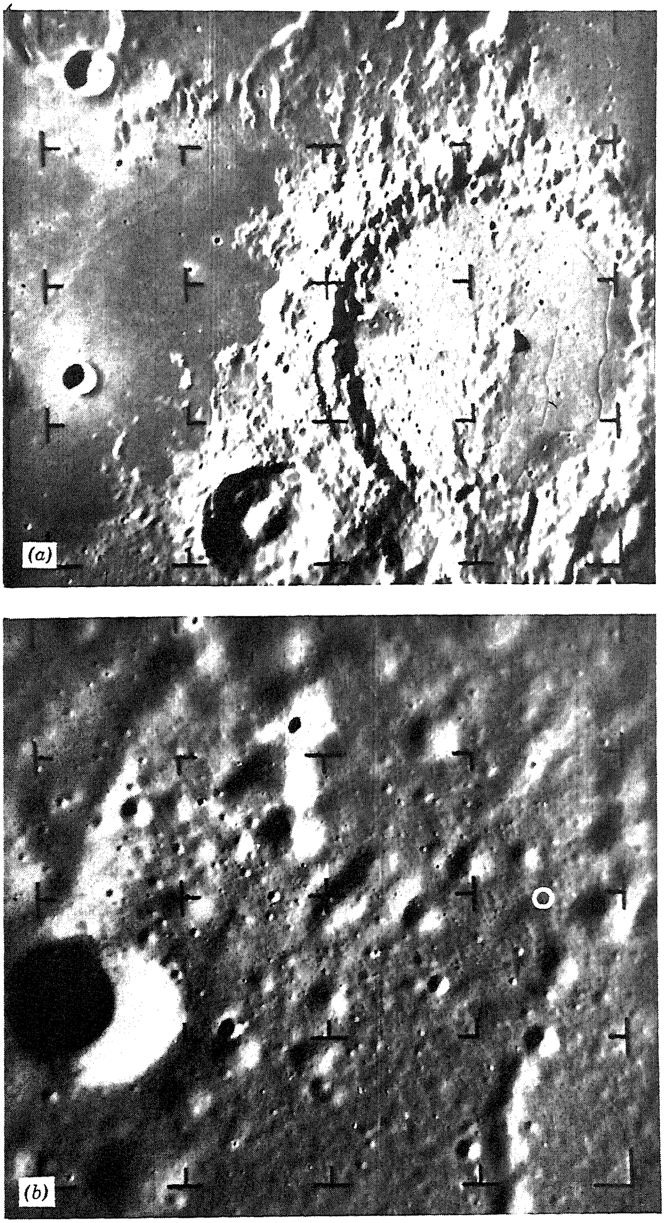


Figure 6-8 (*a*) The crater Alphonsus and the edge of Mare Nubium to the left. Ranger IX struck at the head of a rill marked by the circle. Note the "dark-haloed" craters about the periphery of the floor of Alphonsus. Taken 265 miles above the surface. (*b*) The white circle again marks impact point of Ranger IX and the rill is now seen as a string of collapse craters. Taken 12.2 miles above the surface, 8.09 seconds before impact. (NASA Jet Propulsion Laboratory)

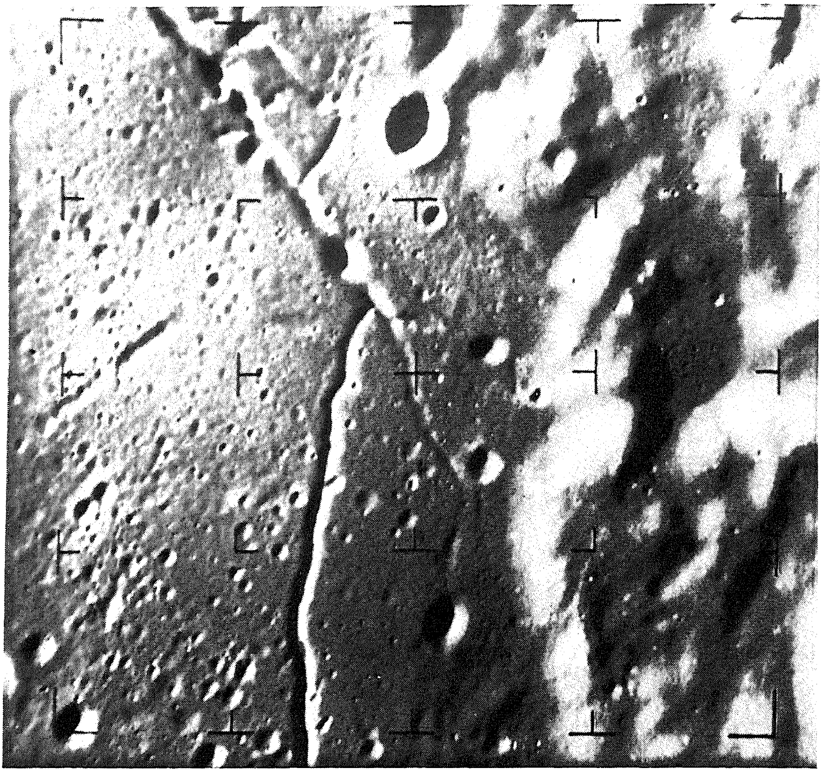


Figure 6-9 A closer view of one of the "dark-haloed" craters taken 107 miles above the surface by Ranger IX. Note the "lava lake" in the smooth walls of the crater Alphonsus. (NASA Jet Propulsion Laboratory)

makes it clear that the dark area immediately surrounding the crater at the bottom of the photograph is smoother than the rest of the surface which is fairly broken up with small craters. It seems obvious that material, perhaps volcanic ash, ejected from this crater has filled in the older craters and even part of the rill which this crater straddles.

Like the central peak, the walls of the crater Alphonsus appear to be smooth, perhaps as the result of the eroding bombardment by meteorites and solar protons and electrons (Figure 6-9). Within the walls there appears to be a high-level "lava lake" with a surface not unlike the floor of Alphonsus.

Another lunar surface feature resembling terrestrial features is the *graben*. A graben is the result of two parallel earthquake faults which

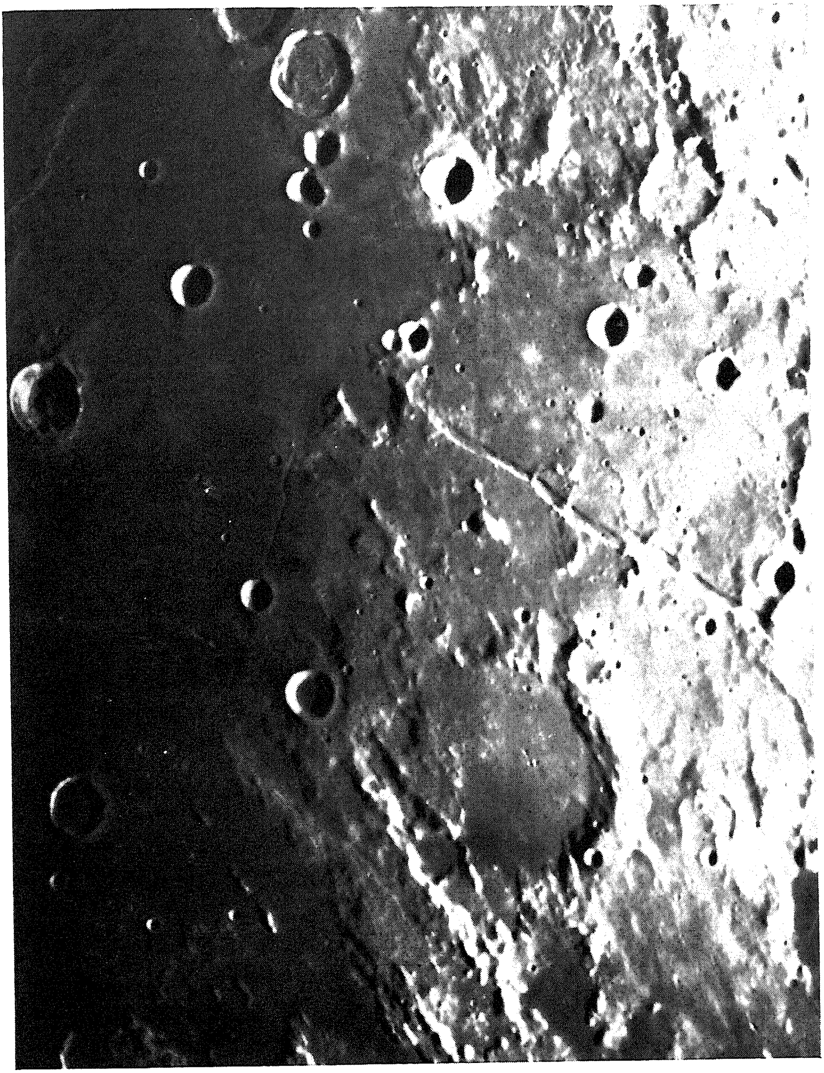


Figure 6-10 A photograph, taken with the 120-in. telescope at Lick Observatory, of a graben extending from the "highlands" into Mare Tranquilitatis. (Lick Observatory)

separate slightly. This separation permits the surface between these two faults to sink. The valley so formed has fairly steep walls, which on the Earth are eroded, usually by water. The region of the moon photographed in Figure 6-10 is the western edge of Mare Tranquilitatis; the large crater toward the bottom is Julius Caesar. Just north of this crater is a graben extending from the right-hand edge toward the center. It would appear as if this graben has been displaced laterally by another fault line intersecting the graben near its center. This photograph was made near the terminator, so the shadows are long and Mare Tranquilitatis is quite dark.

The surface of the moon which presents such a wide variety of features also presents a wide variation in temperature. The hottest parts of the sunlit side reach a temperature as high as 215°F or more. The coldest parts of the nighttime side must have a temperature less than —240°F. Studies of the temperature of the lunar surface have been conducted not only in infrared light but with radio waves as well. It has become quite clear that the surface of the moon does not have a uniform temperature: some regions are warmer than others. Nor does the surface exhibit a uniform change in temperature when the sun light ceases to shine on it.

During a lunar eclipse, when the moon moves into the shadow of the Earth, the temperature of the moon's surface drops. Generally the drop in temperature is quite rapid, amounting to more than 200°F in an hour! Such a sudden drop in temperature is a clear indication that the heating effect of the lunar surface by the sun is not effective to a depth of more than a centimeter or so. The surface is a very good insulator of heat.

There are, however, exceptions to this sudden drop in temperature during a lunar eclipse. The floors of craters which form systems of rays generally hold their heat longer. For example, the crater Tycho, which has an extensive system of rays, holds its heat much longer, indicating that its surface is different from the rest of the moon. Perhaps its surface is freshly (relatively speaking) exposed subsurface rock. Radar signals reflect from the floor of Tycho better than from the surface in general. Again, this indicates a more solid surface than would be presented by a lava flow.

Nearly all of our knowledge of the lunar surface has been restricted to the side nearest the Earth. Space probes launched by the Russians in 1965 have transmitted photographs of the far side of the moon. These photographs indicate that in contrast to the near side, the far side of the moon does not have such extensive large dark areas similar to the maria. The far side is riddled with craters, more in the nature of the highlands of the south polar region of the near side.

With the advent of actual travel to the surface of the moon, interest in our nearby neighbor has increased. New studies have been initiated at a number of observatories. Amateurs have been asked to help by observing the moon in the hopes of recording local color changes which might indicate volcanic activity of some kind. It is hoped that studies of the lunar surface will help answer questions not only about its past, but also about the origin and evolution of the entire solar system. Because of the lack of atmosphere and water, the moon should abound with records of the past.

The Moon's Motion

We have mentioned that the moon revolves about the Earth, but its motion needs to be described in greater detail. Although the moon's diurnal motion is westward, it actually revolves in its orbit from west to east (see page 96). Its orbit is elliptical, as are the orbits of all celestial bodies. The point of its closest approach to the Earth is called *perigee,* and the point farthest from the Earth is called *apogee.*

Our evidence for the moon's eastward motion lies in the fact that it sets on the average 49 minutes later every time it sinks below the western horizon. If the stars are used as reference, the moon has a period of 27⅓ days, during which time it completes a circuit of 360° in the sky. But at the same time the Earth has continued its motion about the sun; thus the moon needs more than 2 additional days, 29½ days in all, to complete its period of phases from new moon through full and back to new moon again. The difference is seen in Figure 6-11. One complete revolution of the moon with respect to the stars (the 27⅓ days) is called a *sidereal period,* while one complete revolution with respect to the line joining the sun and the Earth (the 29½ days) is called a *synodic period.* This latter is the basis for the month.

As has been indicated, the period of rotation of the moon is equal to the period of revolution. Furthermore, since both of these motions are in the same direction, the moon keeps one side toward the Earth. But its rotation and its revolution are not completely synchronized. Since the moon revolves about the Earth in an elliptical orbit, its motion is not steady; it travels fastest when at perigee. But its rotation is steady, so that when the moon changes its speed in its orbit, but not its rate of rotation, its orientation with respect to the Earth changes slightly. The result is that we see slightly more than 50% of the near face of the moon. This effect is called *libration.* As a

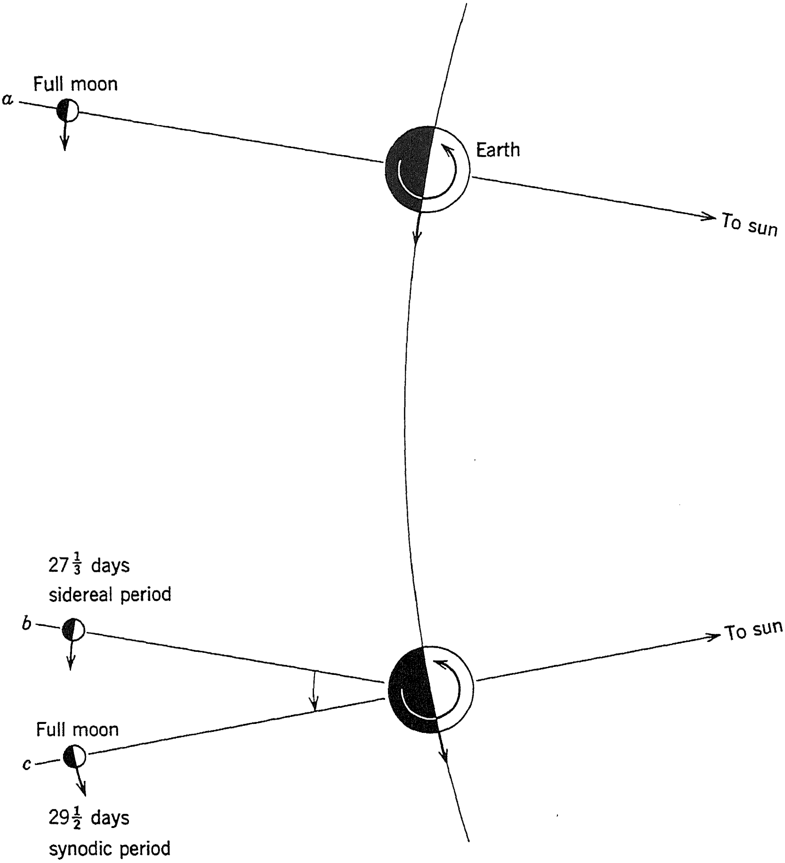


Figure 6-11 The sidereal period of the moon, 27⅓ days, is the time required for the moon to make one complete revolution of 360° with respect to the stars. The synodic period, 29½ days, is the time required for the moon to revolve from one phase to that same phase again, such as from full moon to full moon.

result of libration we see the moon from slightly different directions. This amounts to a stereoscopic observation which permits us to make detailed studies of lunar elevations. The moon is being mapped with contour lines to indicate elevations with respect to an arbitrarily chosen "sea level."

The plane of the moon's orbit is coincident neither with the plane of the Earth's orbit (the ecliptic) nor with the plane of the Earth's equator. It is closer to the ecliptic, however, from which it is tipped by some 5°.

Eclipses

It is because the moon's orbit is not coincident with the ecliptic that the moon does not go into the Earth's shadow every full moon to give us a *lunar eclipse;* nor does the new moon pass in front of the sun every month to give us a *solar eclipse.* Eclipses can occur only during two seasons of the year.

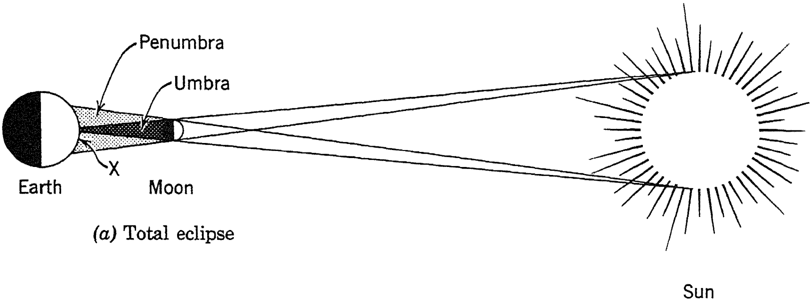
In order for the moon to pass in front of the sun they both must be in the same part of the sky. This means that the moon must be on or very near the ecliptic because, by definition, the sun is always on the eclipitic. The two points in the sky where the moon's path crosses the ecliptic are called the *lunar nodes:* at the ascending node the moon crosses the ecliptic going north, and at the descending node the moon crosses it going south. The moon passes through each node once a month, but the sun passes through each of the nodes only once a year. Thus, a *solar eclipse* can occur only during those two seasons of the year when the sun is at or near one of the nodes during the new moon.

In order for the Earth to cast its shadow on the moon, a *lunar eclipse,* the moon must be at one of its nodes and the sun at the other. Again this can happen only during two seasons of the year. The nodes, however, are not fixed in the sky, for the moon's orbit is precessing much as is the Earth's axis.

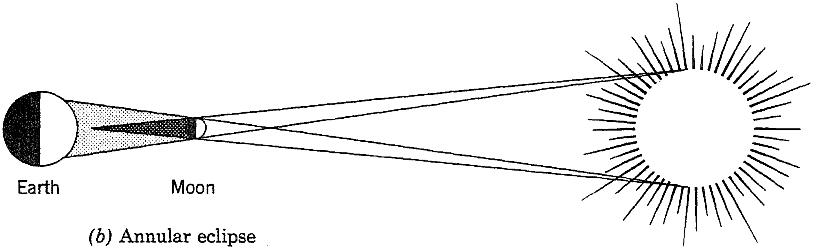
The number of eclipses (lunar and solar) each year varies. It cannot be less than two or more than seven, although the usual number is four. Solar eclipses are more common than lunar eclipses, but since a lunar eclipse can be seen from nearly half of the Earth's surface, more people have seen a lunar eclipse than have seen a solar eclipse.

A total solar eclipse is a beautiful sight. It occurs when the disk of the moon completely covers the disk of the sun (Figure 9-16). But a total solar eclipse is seen only over a small area of the Earth; over a larger area it appears as a partial eclipse. The reason for this is apparent from Figure 6-12*a.* (It has been necessary to draw the sizes and distances of the sun, moon, and the Earth out of proportion to each other.)

In the black region, the sun is completely hidden from view by the moon. This part of the shadow is called the *umbra,* meaning shade or cover (see our word "umbrella"). In the part of the shadow that is gray the sun is not completely hidden. For example, a person standing on the Earth at *X* would be able to see part of the sun. Where this portion of the shadow, called the *penumbra,* strikes the



(*a*) Total eclipse



(*b*) Annular eclipse

Figure 6-12 (*a*) The umbra of the moon just reaches the Earth when the moon is near perigee. (*b*) When the moon is near apogee the umbra does not reach the Earth, and so a total eclipse cannot occur, only an annular eclipse.

Earth, a partial eclipse may be seen. If an observer is near the umbra, a large portion of the sun's disk is covered by the moon, but if he is near the outer edge of the penumbra, only a small section of the sun will be obscured from view. This small section is a silhouette of part of the moon.

The umbra is actually a cone, and it is sheer coincidence that the length of this cone is very nearly the same as the distance from the Earth to the moon. When the moon is near perigee and the Earth near aphelion, the cone of the umbra projects a dark circular shadow about 167 miles in diameter on the face of the Earth. When the moon is near apogee, however, the tip of the cone does not even reach the Earth (Figure 6-12*b*). The resulting visible effect makes the disk of the moon appear slightly smaller than the disk of the sun. Consequently the moon does not cover the sun completely. The ring of the sun that overlaps the moon's disk gives rise to the term *annular eclipse*.

Since the umbra follows the moon in its motion and is projected on a nearly spherical rotating Earth, it traces a rather complicated

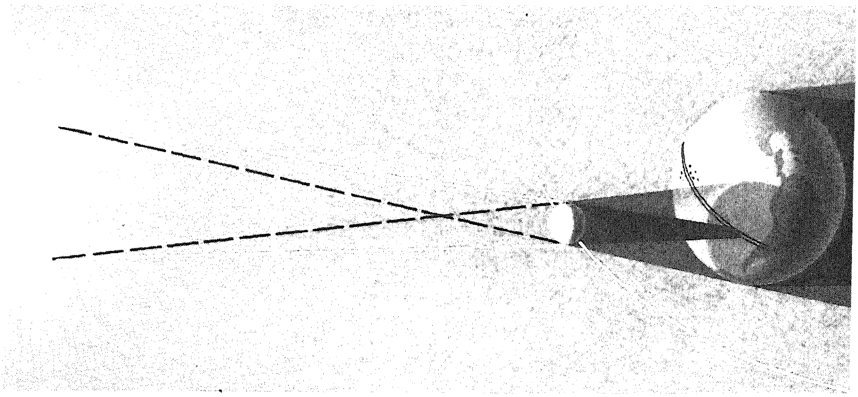


Figure 6-13 As the moon revolves in its orbit and the Earth rotates on its axis, the umbra traces the path of totality across the surface of the Earth. (Courtesy of *Sky and Telescope*)

path, called the *path of totality,* on the Earth's surface. Anyone located in this path may see a total eclipse that can at best last 7 minutes, 40 seconds. The penumbra may stretch for a distance of several thousand miles on either side of the path of totality. The path of totality for the eclipse of October 12, 1958, is seen in Figure 6-13.

At the time of totality the outer regions of the sun become visible. (Because they are fainter than the blue sky they cannot normally be seen.) The pearly white area that extends farthest from the sun is called the *corona.* Very close to the surface of the sun is a region called the *chromosphere* because it is reddish. Prominences that appear to be gases erupting from the sun may be seen standing out from its surface. The nature of these is discussed in Chapter 9.

A total lunar eclipse, although not as spectacular as a total solar eclipse, is a sight long remembered. Even though the moon's disk may be entirely covered by the umbra of the Earth's shadow it is never completely black. The Earth's atmosphere acts like a lens and focuses some of the sunlight on to the moon. Since the red light can travel through a long path in the atmosphere better than the blue, the moon takes on a deep reddish or copper color. The blue light has been scattered out into the surrounding atmosphere.

Eclipses, both lunar and solar, have been beneficial to the historian and archeologist as well as to the astronomer. In the records of ancient peoples there are many references that not only enable the historian to date an event, but also give the astronomer observations extending over many centuries of time. With the help of historic eclipses he has been able to describe the motion of the moon with greater accuracy.

Tides

Not only through eclipses does the moon affect us here on the Earth; the moon is also the main cause of the ocean tides. These tides result from the fact that the gravitational pull of any body, including the moon, decreases with increasing distance from that body. If there were no moon (or sun) the water layer over the Earth would be uniform as shown in Figure 6-14*a*. When the moon is put back in place (shown by the arrow in Figure 6-14*b*) the Earth responds to the additional gravitational force. Those portions of the Earth nearest the moon are subjected to a stronger gravitational pull, and thus the top of the ocean near the moon is pulled closer to the moon than is the bottom of the ocean. On the far side of the Earth, the bottom of the ocean, being nearer the moon than the top, is pulled closer to the moon. Thus the ocean bulges on two sides: the side facing the moon and the side directly opposite the moon. It is in these regions that we experience high tides. Halfway

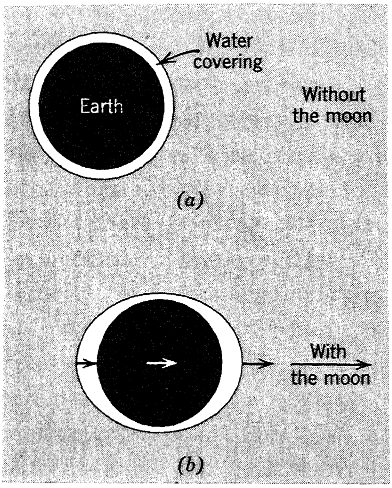


Figure 6-14 (*a*) The layer of water about the Earth in the absence of the moon (and the sun) would be uniform about the equator. (*b*) The moon is there, however, and pulls the Earth towards it. Those parts of Earth closer to the moon are pulled with a stronger force.

between the high tides are the low tides, the regions which supply the water for the high tides.

The net effect of this, and all tidal forces, is that the moon tries to pull the Earth apart but fortunately never succeeds. The Earth tries the same with the moon, and were the moon closer the Earth might succeed.

The Earth rotates underneath the tidal bulges to give us two high and two low tides every day. During one rotation of the Earth, however, the moon has moved eastward in its orbit; consequently the Earth must rotate farther so that the cycle of tides may be completed. Thus there are two high and two low tides every 24 hours, 49 minutes.

The variation in the level of the water differs with locality. Tides of 3 to 6 ft are fairly common on the shores of the oceans, but in some places, because of funneling, the water level may differ by 50 ft between high and low tide. In such places (for example, the Bay of Fundy, which is the gulf between Nova Scotia and New Brunswick) special harbors must be built.

It is bays such as these and shallow waters such as the Bering Strait that cause friction between the tides and the Earth. As a consequence the tidal bulge is dragged by the Earth, which rotates faster than the moon revolves. Hence the tidal bulge does not line up exactly with the moon, but precedes it.

The sun, too, influences our tides, but since it is so much farther away than the moon its effect is not as great. However, at new or full moon the gravitational pulls of the sun and the moon complement each other to give us the *spring tides* (Figure 6-15*a*). During the spring tides the high tides are higher and the low tides are lower than usual. The other extreme, called the *neap tides*, occurs when the moon is at first- or third-quarter (Figure 6-15*b*). During the neap tides the lines joining the sun and the moon with the Earth form a right angle. The moon attracts a high tide from the side of the Earth directly beneath and the opposite side, whereas the attraction of the sun operates to leave only a low tide in these same places. The result is that the high tides are not as high, or the low tides as low as the average.

Since the Earth is not a rigid body it, too, gives slightly with the tides. Although small, this movement has been measured—evidence that having your feet on "solid ground" means something quite different, for that ground is moving up and down every 12 hours, 25 minutes.

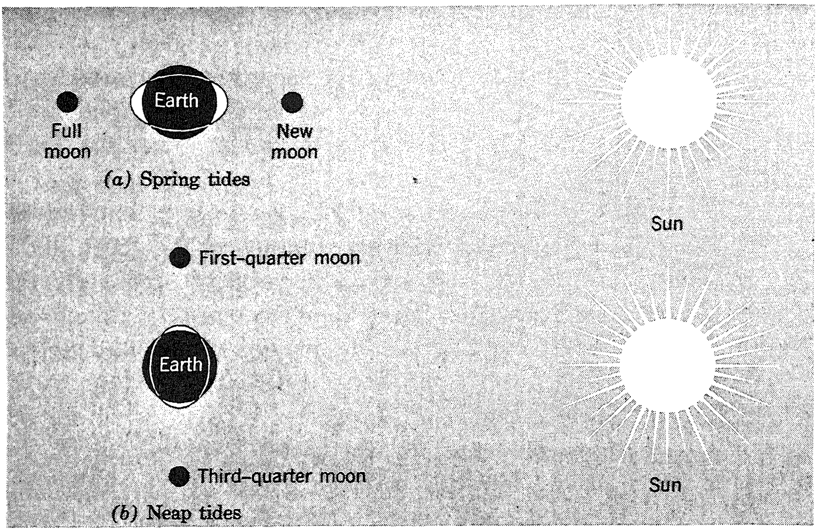


Figure 6-15 (*a*) Spring tides occur when the tidal forces of the moon and the sun work together to produce the high and low tides. (*b*) Neap tides occur when the high tides caused by the sun are superimposed on the low tides caused by the moon.

Other Satellites

Many of the characteristics of our natural satellite can be generalized to apply to satellites of the other planets. Of the nine planets, only Mercury, Venus, and Pluto do not, to our knowledge, have satellites. Those planets that do will be considered next, with emphasis on the unusual satellites.

After the Earth the first planet from the sun to have satellites is Mars with two, Phobos (the inner one) and Deimos. They are small, and their dimness indicates that they are no more than 10 miles in diameter; Deimos is perhaps only 5 miles in diameter.

Phobos is only 5,820 miles from Mars, so close that its period of revolution is only 7 hours, 39 minutes. Since the period of rotation of Mars is 24 hours, 36 minutes Phobos makes a little more than two complete circuits around the planet while the latter is making one complete rotation. This means that Phobos, whose synodic period is 11 hours, 7 minutes, not only rises in the west and sets in the east but may rise twice in one night.

Deimos has a sidereal period of 30 hours, 18 minutes, only a little longer than Mars' period of rotation. Consequently it rises in the ast and progresses very slowly toward the western horizon, where sets about 2⅔ days later. During this long stay in the sky above ny one location it will have gone through its phases twice. But it doubtful that its phases can be seen. The only indication would e a change in brightness from new to full and back to new phase gain.

Jupiter, because of its gigantic size, has more satellites than the un has planets. The first four to be discovered were discovered by alileo in 1609. The fifth satellite was discovered in 1892 and has smaller orbit than the innermost Galilean satellite. The twelfth satellite was not discovered until 1951 by the late Seth B. Nicholson at he Mount Wilson Observatory.

The four satellites discovered by Galileo are called the Galilean atellites and are visible with a modest telescope. By watching them rom night to night an interested observer can follow the motions f these largest Jovian satellites. Two of them are larger than our moon and one, with a diameter of 3,080 miles, is only slightly smaller than the planet Mercury (3,100 miles in diameter).

Since the four Galilean satellites revolve in planes that are very nearly parallel to the plane of Jupiter's equator, they not only go behind Jupiter during part of their cycle but also pass in front of it. During a transit of one of these satellites its shadow is visible on the surface of Jupiter (Figure 6-16).

Of the nine satellites of Saturn (we do not count the tiny particles that compose its rings), Titan and Io are the most interesting because of all the satellites in the solar system, they alone have an atmosphere. This atmosphere, as might be expected, contains methane gas.

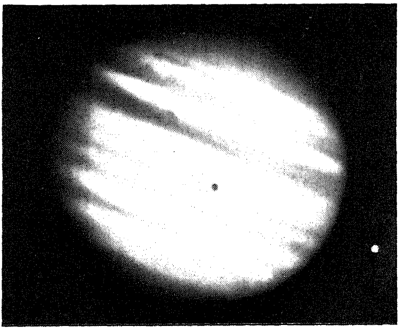


Figure 6-16 One of the Galilean satellites casting its shadow on the cloud covering of Jupiter. (Mount Wilson and Palomar Observatories)

Uranus, whose axis of rotation is nearly parallel to the plane of its orbit, has five satellites revolving about it. Their planes of revolution are nearly parallel to the plane of Uranus' equator. Since these planes are more or less fixed in space, we on the Earth can at certain times (as in 1944) see the satellites moving in circles about the planet; that is, at certain times our line of sight is nearly perpendicular to the planes of their orbits. By the time Uranus has moved in its orbit one-quarter the way around the sun (by 1967), the satellite orbits will be seen edge on and the satellites will appear to oscillate in straight lines.

Neptune has two satellites, one of which when compared to our moon illustrates the dependency of the period of revolution on the mass of the central body. Neptune is 14.5 times as massive as the Earth. One of its satellites, Triton, which has an orbital radius of 220,000 miles (nearly the same as our moon's orbital radius of 239,000 miles), has a period of revolution of only 6 days as compared with our moon's 27⅓ days.

The sun rotates, the planets all revolve and rotate, and most of the satellites revolve and rotate all in the same direction. Triton, however, like the four outermost satellites of Jupiter and the outermost satellite of Saturn, moves in retrograde motion, that is, each of these satellites revolves about its mother planet in a direction contrary to the principal motions of the solar system. It may be presumed that these satellites were captured after their planets were formed.

Manmade Satellites

The subject of satellites is more inclusive now than ever before, for on October 4, 1957, the Russians sent the first manmade satellite, Sputnik I, into an orbit about the Earth. The question is often asked: "How does a satellite stay up there?" The answer to this also answers the questions: "Why doesn't the Earth fall into the sun?" and "Why does a rock whirled at the end of a string keep the string taut?"

As was stated in the discussion of Newton's second law (see page 100), if any object in motion has no force acting on it that object will move in a straight line and with a constant speed. If a force should now act on this object, then according to Newton's second law of motion the object must accelerate. The amount of acceleration depends not only on the force applied but also on the mass of the object. The direction of acceleration is the same as the direction in which the force is acting. If the force is applied in a direction that is perpendicular to the direction of travel, the direction of travel will change accordingly. If this force is constant and always acts perpendicular to the direction of travel, the object will move in a circular path or orbit (Figure 5-3).

Any force that causes an object traveling in a straight line to be diverted into a curved path is called a *centripetal force.* The centripetal force acting on the whirling rock is the tension in the string. The centripetal force that keeps the moon or artificial satellite in its orbit is supplied by the mutual gravitational pull of the Earth and the satellite. This force of gravity is just strong enough to continually divert the satellite from its tendency to travel in a straight line, into an orbit that may be circular or elliptical.

The fact that the force of gravity is exactly enough to keep the satellite traveling in an orbit has consequences for any object that rides in the satellite. For example, suppose that a satellite were cut into two pieces, one half moving along just outside the original orbit and the other half moving just inside. The two satellites would travel along side by side for some time, neither one falling to the Earth any more than the other. Of course, eventually the outer satellite would lag behind, for its orbit is slightly larger and thus its velocity slightly less and its period slightly longer (Kepler's third law).

If, now, rather than cutting a satellite in two, we place one satellite inside the other, they would then have the same orbit and they would travel around together, neither one tending to fall toward the Earth any more than the other. When man rides inside a satellite, he is the inside satellite; thus he is in an orbit and feels weightless. The lack of apparent weight results from the fact that the entire force of gravity is used up in keeping the satellite and man in an orbit. In the same manner a man diving off a high diving board is weightless, since the entire force of gravity is used to accelerate him as he falls.

The motions of artificial satellites, since they are not powered, are described by Kepler's three laws just as natural satellites and planets are. The length of each period depends on the mean radius of its orbit as well as the mass of the Earth. The mass of an artificial satellite is so small when compared with that of the Earth that its orbit may be considered to be independent of its mass. The mean radius and the shape of the orbit depend on the height to which the final-stage rocket finally propels the spacecraft, the final speed, and its direction of travel.

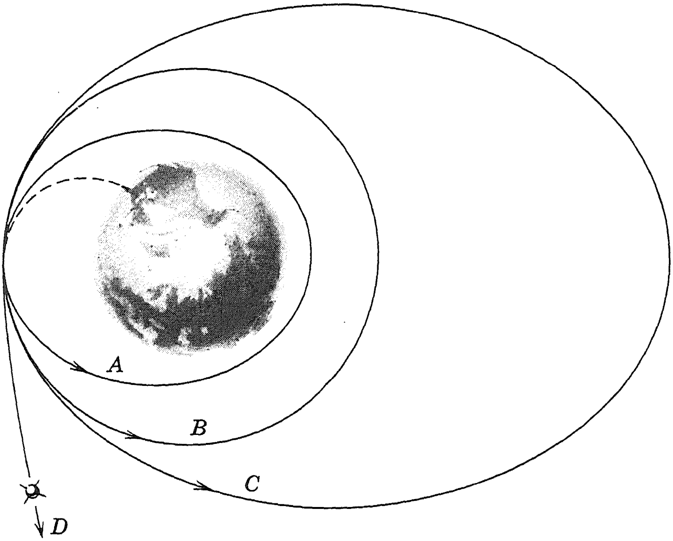


Figure 6-17 Three possible orbits for a satellite. The velocity with which the satellite was placed in the orbit helps determine the type of orbit.

If the final speed and direction are just right, the satellite will travel in a circular orbit concentric with the Earth (Orbit *B*, Figure 6-17). If, however, the final velocity is less than that for a circular orbit, and its direction of travel is parallel to the local horizon, the altitude at which the final velocity was achieved will become apogee, and it will travel closer to the Earth to establish perigee (Orbit *A*). If the final velocity is greater than that for a circular orbit, it will establish this altitude as perigee and travel farther from the Earth to apogee (Orbit *C*). The velocity for a circular orbit whose radius is 4,400 miles (that is, 400 miles above the surface of the Earth) is about 18,000 miles per hour or about 5 miles per second. With orbits of increasing radii the velocity is less; the velocity of the moon in its orbit is 0.66 miles per second or about 2,400 miles per hour.

If, however, the speed of the spacecraft just equals the escape velocity of the Earth, its orbit would assume the shape of a *parabola* (Figure 6-18). The shape of the orbit would become a *hyperbola* if the speed of the craft exceeded the escape velocity. The ellipse is a closed figure, but both the hyperbola and the parabola are open curves, that is, the spacecraft would not return to Earth. It would, however, assume an elliptical orbit about the sun, unless, of course,

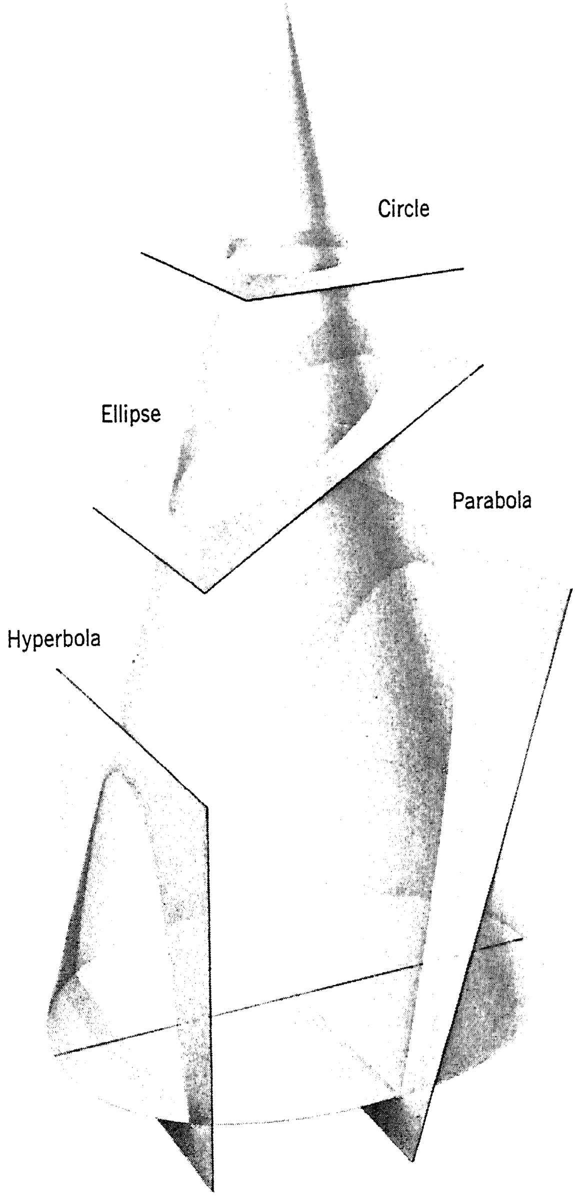


Figure 6-18 The intersection of a plane and a cone is called a *conic section.* If that plane is perpendicular to the axis of the cone the conic section is a *circle;* if that plane is oblique to the axis of the cone, yet cuts every element of the cone, the intersection is an ellipse. If that plane is parallel to one element, the intersection is a *parabola;* if the plane is parallel to the axis of the cone the intersection is a *hyperbola.*

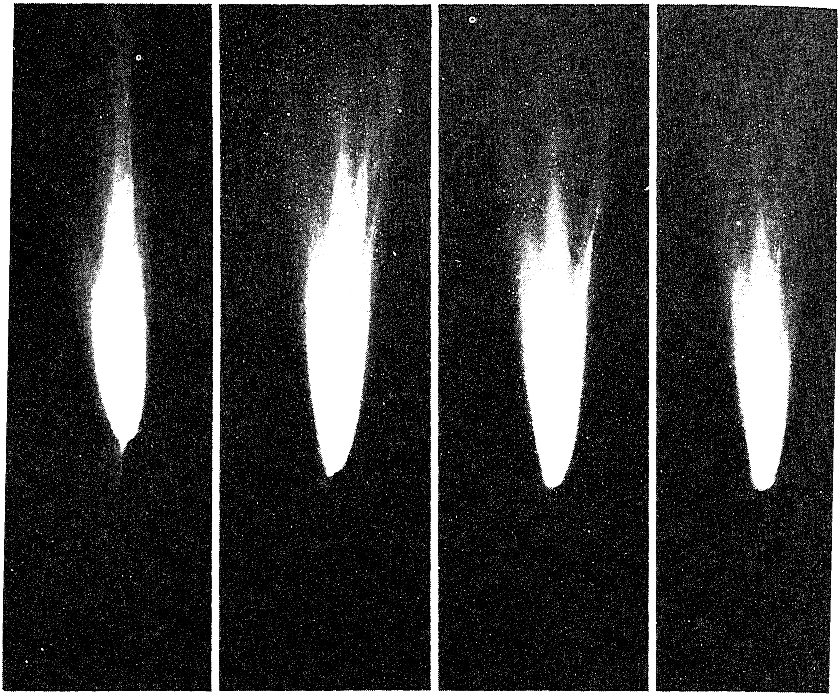
its speed should reach or exceed the escape velocity for the sun at this distance from that central body.

Basic Vocabulary for Subsequent Reading

Annular eclipse
Apogee
Lunar eclipse
Neap tides
Occultation
Partial eclipse
Penumbra
Perigee
Sidereal period
Solar eclipse
Spring tides
Synodic period
Total eclipse
Umbra

For Further Reading

Alter, Dinsmore, *Pictorial Guide to the Moon,* Thomas Y. Crowell Co., New York, 1963.
Baldwin, Ralph B., *A Fundamental Survey of the Moon,* McGraw-Hill paperback, 1965.
Firsoff, V. A., *Strange World of the Moon,* Basic Books, New York, 1959.
Mitchell, Samuel A., *Eclipses of the Sun,* Columbia University Press, New York, 1951.
Whipple, Fred L., *Earth, Moon, and Planets,* Harvard University Press, Cambridge, Mass., 1963.
Greenacre, James A., "A Recent Observation of Lunar Color Phenomenon," *Sky and Telescope,* p. 316, December 1963.
Kuiper, G. P., "Lunar Results from Ranger 7 to 9," *Sky and Telescope,* p. 293, May 1965.
Kuiper, G. P., "The Lunar and Planetary Laboratory," *Sky and Telescope,* part I, p. 4, January 1964; part II, p. 88, February 1964.
O'Keefe, John A., "Lunar Questions Discussed at Pasadena," *Sky and Telescope,* p. 10, January 1966.
Saari, John M. and Richard W. Shorthill, "Hot Spots on the Moon," *Sky and Telescope,* p. 327, June 1966.
Schumeier, H. M., et al., "The Ranger Missions to the Moon," *Scientific American,* p. 52, January 1966.
Shoemaker, Eugene M., "The Geology of the Moon," *Scientific American,* p. 38, December 1964.
Wildey, Robert L., "Measuring the Shape of the Moon," *Sky and Telescope,* p. 147, March 1966.
"Icy-Conglomerate Moon," *Sky and Telescope,* p. 143, March 1965.



April 27 April 29 April 30 May 1

1957

Five views of Comet Arend-Roland taken with the 48-in. Schmidt telescope on Mount Palomar. (Mount Wilson and Palomar Observatories)

Chapter 7 *Minor Planets, Meteorites, and Comets*

Man leaves no stone unturned in his eternal search for system and law. This search has led from the Pythagorean numbers games to the bases of modern mathematics; and numbers games are still being played, with results that are sometimes as fruitful.

Take, for instance, the mean distances in astronomical units of the planets from the sun. Beginning with Mercury, we find that these give, naturally enough, a sequence of increasing numbers:

0.387, 0.723, 1.00, 1.52, 5.20, 9.54, 19.2, 30.1, 39.5

An early attempt to impose logic and system upon this sequence is known as *Bode's law*. No law was ever less legitimate; it is merely a scheme for systematically duplicating an approximate progression of planetary distances from the sun. If we start with 0.0, 0.3, 0.6, etc., doubling each number until we reach 38.4, we obtain the following sequence simply by adding 0.4 to each of the resulting numbers:

0.4, 0.7, 1.0, 1.6, 2.8, 5.2, 10.0, 19.6, 38.8

This sequence roughly approximates the natural one, except that between the numbers 1.6 and 5.2, which represent the distances of Mars and Jupiter respectively, there is a 2.8. Also, the number 30.1, representing the distance of Neptune, is missing, and the number 38.8 differs considerably from the actual distance of Pluto.

All this is quite meaningless. It is a game of numbers.

It is possible, however, to derive a sequence that is more descriptive of the physical facts. Let us divide the mean radius of each planet's orbit, starting with Venus, by that of the planet just inside its orbit. The result indicates how many times farther from the sun a given planet is than its neighbor toward the sun. For example, 0.723 divided by 0.387 equals 1.9, which tells us that Venus is 1.9 times as far from the sun as Mercury. Doing this for all the planets, from the nearest to the farthest, gives us the sequence:

1.9, 1.4, 1.5, 3.4, 1.8, 2.0, 1.6, 1.3

All these numbers lie between 1.3 and 2.0 except for that single 3.4, obtained by dividing Jupiter's distance from the sun by that of Mars. These two planets, then, are farther apart with respect to their distance from the sun than are any other neighboring planets.

The Minor Planets

If Bode's law is not taken seriously today, it did gain support in 1781 when the discovery of Uranus appeared to verify it; the predicted 19.2 is not far from the observed 19.6. On January 1, 1800, Bode's law acquired still another friend. The Italian astronomer Piazzi discovered a very small planet, later named *Ceres,* whose distance from the sun is the expected 2.8. In March, 1802, a second small planet, *Pallas,* was discovered at about the same distance from the sun—and the race was on. Since that time, thousands of very small *minor planets* (sometimes called *asteroids*—but they are "star-like" only when seen from great distances) have been discovered, most of which revolve about the sun in that gap between Mars and Jupiter.

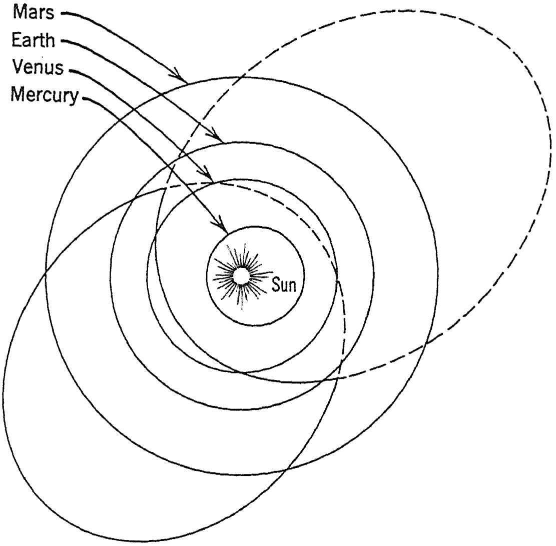


Figure 7-1 The orbits of two minor planets that come inside the orbit of Mars. The dotted portions of the orbits are south of the plane of the ecliptic.

Each minor planet is nothing but a chunk of rock revolving about the sun. These chunks of rock, however, should contain an invaluable record of historical events in the solar system.

The distance of each minor planet from the sun can be found by computing its orbit from observed positions. Most minor planets have orbits that lie between the planets Mars and Jupiter. Some, however, come inside Mars' orbit and others inside the orbits of Venus and even Mercury. As a group, then, these very populous members of our solar system travel over a considerable range. Their orbits are sometimes quite eccentric, that is, elliptical and considerably distorted from a circle. Two of these orbits are shown in Figure 7-1.

Minor planets are discovered most frequently by accident. A photograph of a field of stars is usually a time exposure because the stars are faint; and since a minor planet moves with respect to the stars, it will leave a blurred streak instead of a point as do the stars (Figure 7-2). (The telescope is usually driven and guided to follow the stars.) If the minor planet leaves a long streak it is in all likelihood relatively close to the Earth and is therefore of interest. It is then further observed in order to obtain positions that are required for the calculation

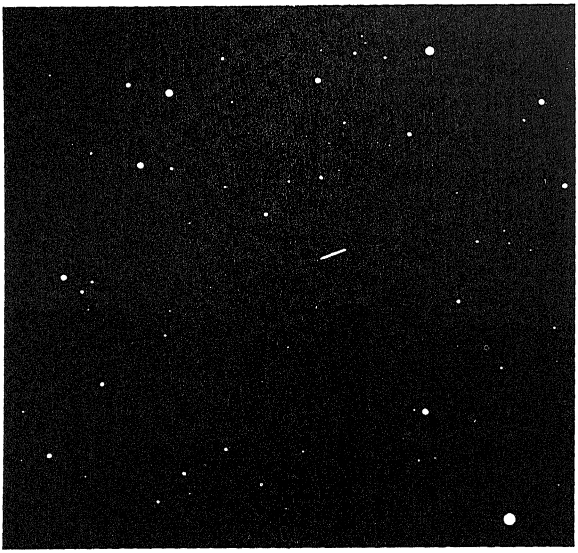


Figure 7-2 The motion of a minor planet causes it to appear as a streak on a time exposure of the stars. (Courtesy of C. Wirtanen, Lick Observatory)

of its orbit. If the minor planet leaves a short streak, however, it is of little interest as it is likely to be far from the Earth and one of the many that may be forgotten.

It may seem "unscientific" that the astronomer should disregard any evidence at all, but it is time-consuming to make observations and compute orbits. There are so many minor planets, so few facilities, and so many pressing unknowns in the universe that he is forced to deal only with those unknowns which he decides are most important.

When the periods of a large number of minor planets are arranged according to their length, we may note an interesting phenomenon (Figure 7-3): there are periods that seem to be avoided. The most conspicuous are those of 6.0, 4.8, and 4.0 years which happen to be ½, ⅖, and ⅓ of the 12-year period of revolution of Jupiter. Other avoided periods are ¼, ⅕, and ⅗ of the period of Jupiter.

This is not a coincidence but a direct result of the perturbing effect massive Jupiter has on these small objects. The missing periods can be translated into mean distances from the sun by Kepler's third law; they represent actual gaps in space that are populated by fewer minor planets than other regions in the "minor planet belt." These

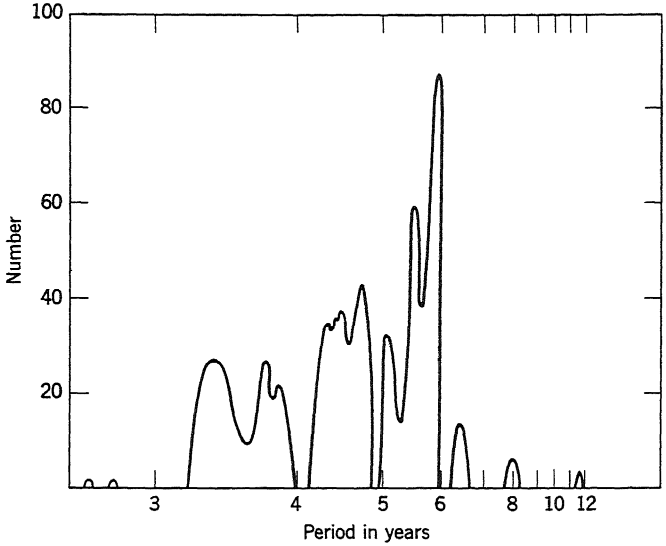


Figure 7-3 **A distribution of minor planets according to their periods indicates that certain periods are rare.**

gaps are similar to the gaps between Saturn's rings which are caused by the perturbing effect of Saturn's innermost satellite.

Since minor planets are so small, it is difficult to estimate the size of any but the very largest. It is possible, however, to make a reasonable estimation by reversing the procedure for calculating the albedo of a body. A minor planet is too small to hold an atmosphere and therefore we can assume its albedo, like that of the moon and Mercury, to be 0.07. Knowing its distance from the sun, we can compute how much light the minor planet receives; and knowing its distance from the Earth and assuming this albedo, we can calculate how large it must be to reflect as much light as it does.

Fortunately, a few of the brightest minor planets have been seen as disks, and so the sizes of these minor planets have been calculated. Ceres is the largest with a diameter of 480 miles; Pallas is next with a diameter of 300 miles; then Vesta, 240 miles; and Juno, 120 miles. Since their diameters have been measured their albedoes can be determined; as might be expected they are quite small. The smallest albedo, 0.06 for Ceres, is about the same as that of the moon. A few exhibit higher albedoes, for example, Juno 0.12 and Vesta 0.26; this indicates that these must have a surface that is different from, and more reflective than, that of the moon and Mercury.

It is not uncommon to find a minor planet whose brightness varies. Since every minor planet is seen by reflected sunlight, the natural assumption is that it must rotate on its axis. Consequently, it presents to the Earth different faces which may have different reflecting qualities; or, more probably, the minor planet itself may be irregular rather than spherical in shape and by rotating may present now a larger and now a smaller cross section to the Earth. This effect has been observed with oblong artificial satellites that tumble, such as Sputnik I and II.

The minor planets, because of their size and number, must be listed as a separate group in the membership of the solar system. Why they are so different remains one of the unknowns of this system, but the given numerical sequences have suggested to some that perhaps there was once a planet between Mars and Jupiter which for some reason broke up into many fragments. If there were such a planet, we would be left to wonder about the cause of the break-up.

Meteorites

Compared to the nine major planets the minor planets are very small, but they are far from being the smallest objects that revolve about the sun. *Meteorites* are not only smaller, but since they collide with or fall on the Earth they have become the first extraterrestrial objects to be touched by man and examined in his laboratory.

Meteorites travel in orbits about the sun at velocities of 10 to 30 miles per second. When one collides with the Earth it must first penetrate our protective atmosphere. During this high-speed penetration, the air in front of the meteorite does not have time to move to the side, so the meteorite acts like a piston, compressing the air in front of it. The heat resulting from this compression causes the meteorite to melt and burn. That is, the energy, both kinetic and potential, which the meteorite had with respect to the Earth is converted into heat. Small meteorites never survive this burning; they filter down to the surface of the Earth as small particles of ash or dust. From the length of time it takes an average meteorite to burn—just a fraction of a second—one can estimate their size to be not much larger than an "o" on this page.

As the meteorite streaks and burns through the upper atmosphere, it may be visible from the earth as a "shooting star" or *meteor*. The term meteor derives from the fact that before the very early part of the nineteenth century meteoric streaks were thought to be a

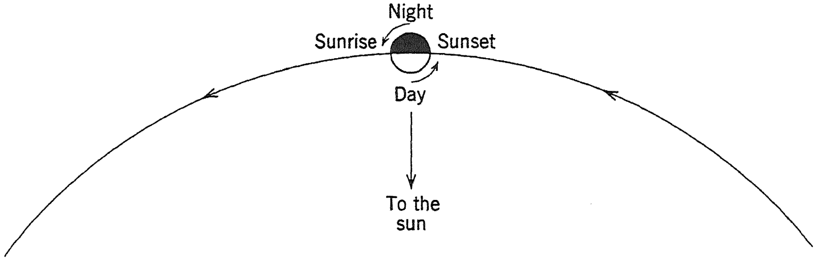


Figure 7-4 The sunrise portion of the Earth leads the way as the Earth revolves about the sun.

strictly atmospheric phenomenon. A patient observer away from city lights may observe 5 to 10 meteors every hour. If he continues watching all night long he will find that the frequency of meteors increases as the night progresses. Not only will the number increase each hour but after midnight they will become, on the average, more brilliant.

We can understand the increase in meteor activity in the early morning hours if we recall how the Earth revolves and rotates. It rotates in the same direction in which it revolves, so the leading side of the Earth is the morning side (Figure 7-4). That portion of the Earth in dawn, is, therefore, "out in front" of the Earth as it travels in its orbit about the sun. As a consequence it is able to pick up more meteors; it runs into them. Thus the early morning meteors have a higher velocity when they strike our atmosphere. The brightness of a meteor depends to a large extent upon its relative velocity, so these meteors are on the average more brilliant than those seen in the evening. An evening meteor overtakes the Earth and therefore has a low relative velocity.

Most meteors are rather faint, but sometimes a very bright one rivals the brightness of the full moon and even casts shadows. Such a bright meteor is generally called a *fireball.* Since the difference between a meteor and a fireball is dependent on brightness, it is difficult to make a definite distinction between the two. In his book *Between the Planets,** Watson calls attention to a very quaint way of distinguishing a meteor from a fireball. He says that, "One astronomer adopted a very practical definition: 'A fireball is a meteor sufficiently bright to make people report it.'" Some fireballs may be so bright as to be seen over the areas of several states.

* Fletcher G. Watson, *Between the Planets*, Cambridge, Harvard University Press, 1956.

At best meteors are difficult to study, for they give no warning when or where they will be seen, and they last only a fraction of a second once they strike our atmosphere. Thus it is hard to set up a planned program for observing and studying them as has been done for other astronomical objects of interest.

The first study of meteors was made in 1798 by two German students who realized that each had seen the same meteor even though they had been separated by several miles on the Earth. In following up this observation with others, they learned how to determine the height of the meteors by the method of parallax (Figure 7-5). Their results showed that meteors appear at a height of about 50 miles.

In the past few decades Harvard College Observatory has made a concerted effort to study meteors. A camera (the Baker-Super-Schmidt) has been designed specifically for their investigation. Not only is it able to observe a large part of the sky and increase the chance of catching an unexpected meteor, but also it has a rotating shutter which interrupts the light path at known time intervals. This makes the path of the meteor appear as a series of short dashes. Since we know the time interval between breaks we need only count the number of dashes to learn how long the meteor has been emitting light.

By setting two such cameras some miles apart on the Earth and pointing them to the same region in the upper atmosphere they are apt to photograph the same meteor, but from different directions (Figure 7-5). Comparing the two pictures enables us to determine the position of the meteor by parallax. Knowing its position at the beginning and end of its trail gives us the distance it has traveled, and dividing this by the time gives us its average velocity.

That velocity may range from about 6 miles per second to about 50 miles per second. The faster meteors become visible at heights of up to 75 miles and generally disappear at a height of about 60 miles; the slower ones appear at about 50 miles and disappear at about 30 or 40 miles. Since the faster meteors are more subject to heating by the atmosphere, they make their appearance higher where the atmosphere is thinner.

Formerly, when only optical telescopes were used, the study of meteors could be carried on only at night. The use of radio telescopes has made possible daytime studies using the radio frequency range of the electromagnetic spectrum. As a meteor streaks through the upper atmosphere it not only becomes very hot but as a result of

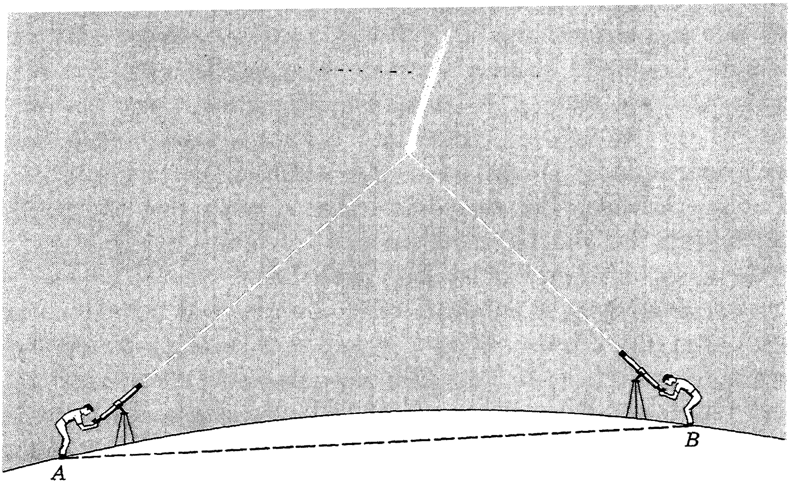


Figure 7-5 The height of the meteor can be determined if it is observed either visually (without a telescope) or photographically (with a telescope) from two well-separated positions on the Earth.

its speed and heat ionizes the atoms in its path. The resulting path of ionized gas reflects radar waves. This gives us a method of observing meteors in the daytime: radar signals can be sent to the upper atmosphere and the echoes or reflections that return can give us the distance and direction of the meteors. The Doppler shift of the reflected signal yields the velocity of the meteor.

By this method, meteor counting can be greatly simplified; the signal, after amplification, is simply shown on an oscilloscope (similar to a television tube) and a photographic record made. Such counts have indicated that generally the maximum number of meteors strikes the Earth's atmosphere at about 6 A.M.

Meteors may help us understand the upper atmosphere but meteorites are of greater interest to astronomers as extraterrestrial objects. Not only are they tangible evidence that the chemical elements on the Earth are no different from those in the solar system in general (which is certainly to be expected), but also they have been helpful in determining the solar system's age. They are, incidentally, a bundle of clues that need to be untangled to help solve the mystery of the origin of the solar system.

The meteorites can be classified into three general groups: iron-nickel, stony, and stony-iron. The iron-nickel meteorites average about 90% iron and about 8% nickel (although these percentages vary considerably) with the balance consisting of other elements. The stony meteorites have quite a different composition; their main chemicals are 36% oxygen, 24% iron, 18% silicon, 14% magnesium, (again, these are averages and there is considerable variation), and a smattering of other elements. The stony-irons form a group that bridges the gap between the other two main groups.

More iron-nickel meteorites have been found than all others put together. But iron meteorites are easily distinguished from terrestrial rocks, even by a farmer tilling the soil, or a hiker or hunter. Iron meteorites have been found which, according to radioactive studies (see Chapter 8), have been on the Earth for as long as 800,000 years. Many of the stony meteorites would have weathered away in that time interval. Of those observed to fall and then recovered, only a little more than 40 have been iron meteorites; more than 600 stony meteorites have been seen to fall and then collected! Nevertheless, all of the big meteorites are significantly iron-nickel and they far outweigh all of the stony meteorites.

The reason for the existence of three main types of meteorites is not really known, but they have fed speculation on a possible relationship between the minor planets, the meteorites, and the "missing planet" between Mars and Jupiter. It is reasoned that perhaps there were two small planets in this region while the solar system was forming. As these planets formed, the heavier elements would have settled to the center as they have done in the Earth. Therefore, the iron-nickel meteorites would have formed the cores of these two smaller planets. The stony meteorites would have composed the mantle and whatever crust existed. The stony-iron meteorites may represent the transition between the core and the mantle.

If these two planets were forming at about the same distance from the sun they may have collided, spreading the smaller fragments over the inner part of the solar system while the larger chunks remained more or less in the same orbit to constitute the minor planets. The perturbations caused by Jupiter, and to a much lesser extent by some of the other planets, would have spread out the minor planets into the belt they now occupy.

There is other evidence that leads us to suspect that the meteorites were formed inside a much larger body. When meteorites are examined in the laboratory it is found that they contain crystals. This in itself is nothing startling, but the crystals are large (Figure 7-6). If a rock contains large crystals the geologist knows that the rock cooled and the crystals formed very slowly; if they are small he knows that the rock cooled and the crystals formed more rapidly. Since the crystals in meteorites are large they may once have been a part of a much larger body that took some time to cool from a molten state.

Detailed studies of the crystals by J. A. Wood of the Enrico Fermi Institute for Nuclear Studies have permitted good estimates of the rate of cooling of the iron meteorites. One group of meteorites studied apparently cooled at the rate of 1°C per 100,000 years; the other group cooled 10 times more slowly. This evidence indicates that two parent bodies are involved: the faster-cooling meteorites came from a body about 60 to 120 miles in diameter, the slower-cooling meteorites from a body about 160 to 200 miles in diameter.

There are suggestions, on the other hand, that meteorites come from the moon. Certainly, if a large meteorite or comet struck the surface of the moon, some of the material would be ejected at velocities exceeding the escape velocity of the moon. Curious objects called *tektites* have been particularly singled out as lunar in origin. Tektites are found in strangely limited areas of the world: in Australia, Java, the Philippines and on the Indochina peninsula. Quite different tektites are found in Czechoslovakia, Africa, Texas, etc. The tektites are unique in their glass-like composition and more notably in their shapes resulting from heating apparently upon entering our atmosphere at velocities less than that which burn meteorites.

It is significant that the shapes of tektites has been duplicated in the laboratory by using pieces of tektite glass mounted in a high-speed wind tunnel (Figure 7-7). But more than that, by a thorough study of surfaces and flow lines, the experimentors, Dean R. Chapman and Howard K. Larson of the Ames Research Center, California, have been able to establish the primary origin of tektites in an atmosphere with a pressure less than 0.0002 times the atmospheric pressure at the surface of the Earth. This observation would eliminate terrestrial volcanoes as the source of tektites. Furthermore, these studies indicate that their velocity upon entering the Earth's atmosphere must have been just slightly higher than 7 miles per second, the escape velocity at the surface of the Earth. There seems little question that tektites indeed originate on the moon, and probably result from the impact of meteorites on the moon's surface.

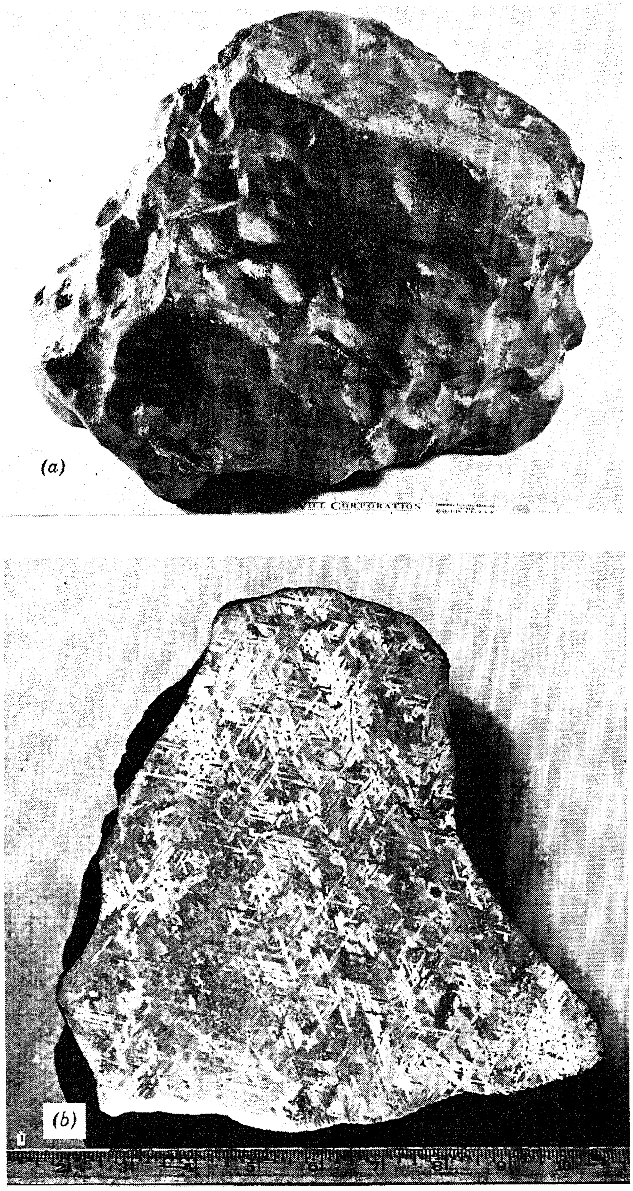
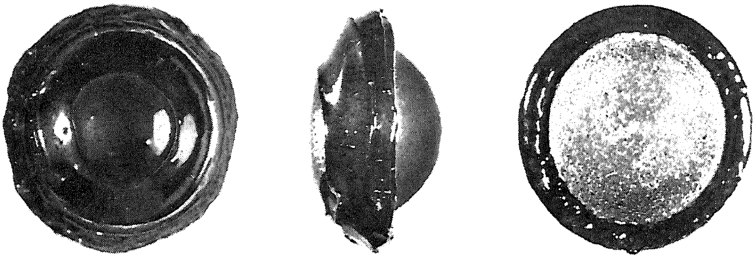
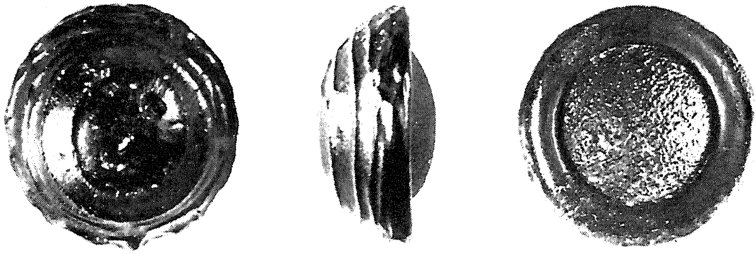


Figure 7-6 (*a*) A typical stony meteor showing on its surface the results of heating by the atmosphere. (*b*) Crystals in an iron-nickel meteor found in Willow Creek, Wyoming. (American Meteorite Museum, Sedona, Arizona)



AERODYNAMIC ABLATION OF TEKTITE GLASS



AUSTRALIAN TEKTITES

Figure 7-7 **Three different pieces of tektite glass which have been placed in a wind tunnel and subjected with an arc jet to simulate entry into the atmosphere. These are compared with three natural tektites found in Australia. (Courtesy Dean R. Chapman, NASA)**

Another class of meteorites has proved every bit as interesting as the iron-nickel meteorites. The stony meteorites have been divided into two subclasses: the *achondrites* and the *chondrites.* The chondrites contain little spherical bodies only about 0.1 cm in diameter, called *chondrules.* Study of these chondrites has opened up new avenues of approach to the age-old problem of how the solar system came into being.

The study of chondrites has proceeded along lines similar to studies of the nickel-iron meteorites. Both objects, after all, are basically crystalline in structure. As such they are the subject of the geochemist. The geochemist studies the growth of crystals, the chemical reactions which lead to the particular formations, and conditions of temperature and pressure which must prevail before a particular crystal or chemical can form. With this background, he can learn a great deal about the history of the different kinds of meteorites. Studies of chondrites have, it appears, led to some enigmas.

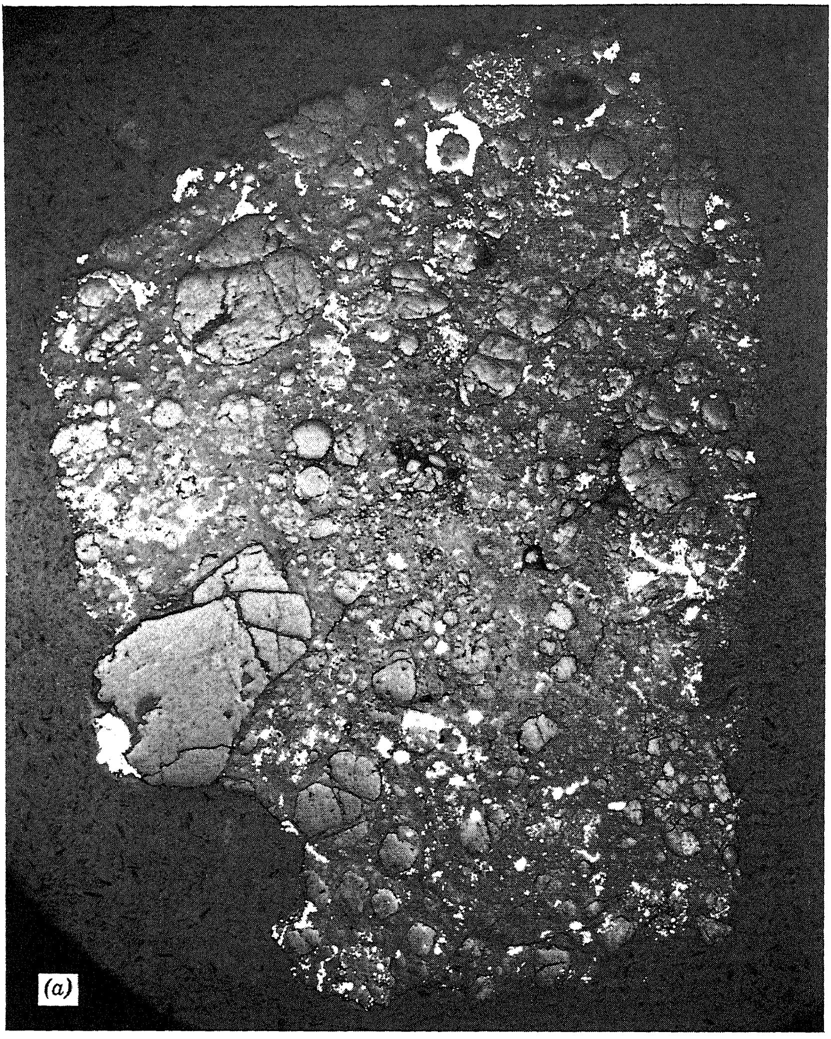


Figure 7-8 (*a*) **Chondrules embedded in the surface of the Bjurböle (Finland) chondrite meteorite.**

The chondrules are held in a matrix composed to a large extent of two minerals found in terrestrial rocks: olivine and pyroxene (Figure 7-8*a*). The matrix also contains other minerals, including glass; the matrix is often rather friable and as such crumbles easily. But the principal distinction between the matrix and the chondrules is their iron content.

This distinction is particularly evident in a class known as *carbona-*

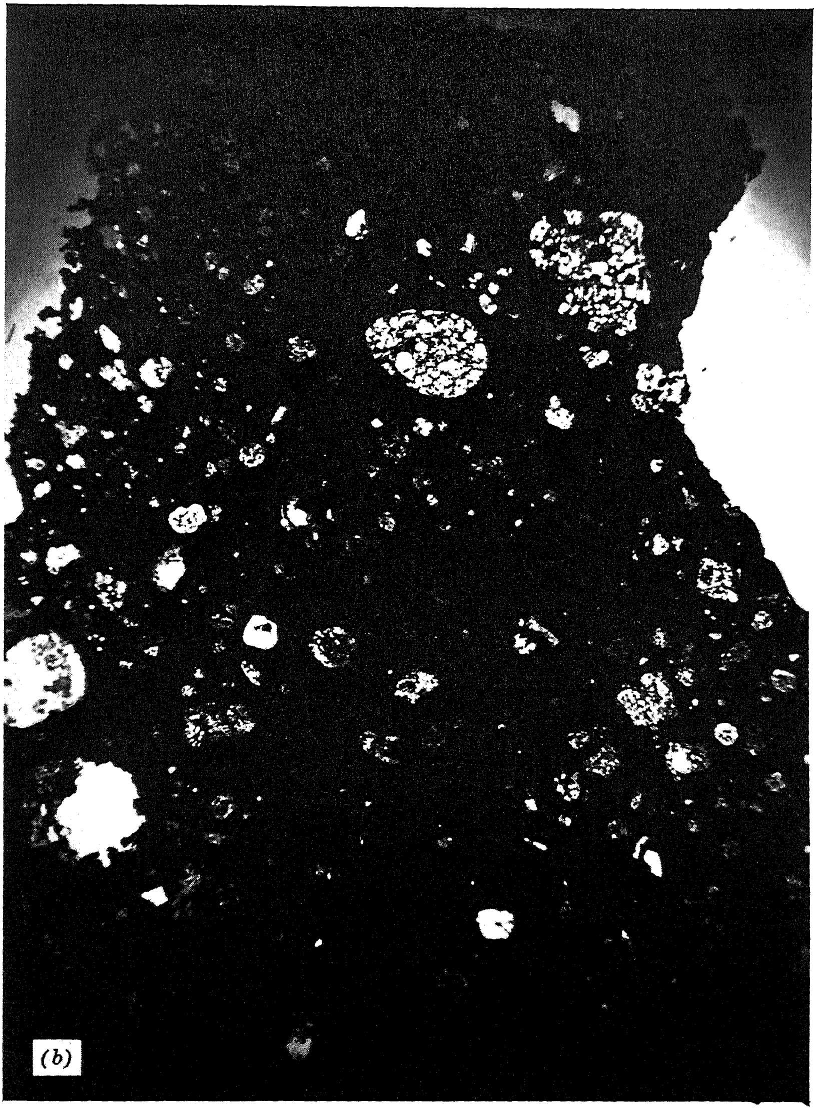


Figure 7-8(*B*) (*b*) A thin section of a carbonaceous chondrite taken from the Murray meteorite which fell in Kentucky. (Courtesy Brian Mason, U.S. National Museum) *ceous chondrites,* because of the black, loosely held matrix (Figure 7-8b). The matrix contains almost no free iron; nearly all of the iron has been combined chemically to form oxides and other compounds. The iron in the chondrules, however, is almost entirely free; less than 1% occurs as iron oxide.

Furthermore, it is found that the chemical composition of the chondrules is almost identical with that of the sun. This cannot be said of any rock on the Earth. During the Earth's history, rocks have been heated and melted so that chemicals have become separated. This has not happened to the chondrules in the carbonaceous chondrites.

Radioactive studies lead to the conclusion that these bodies were formed in the very earliest stages of the solar system—in the original gases from which the planets were made. Their content of free iron further indicates that they were formed during the passage of a shock wave set up by the forming and partially unstable sun, a shock wave which was transmitted through the primordial gases. It is presumed that during the shock wave some of the material would condense into liquid droplets and harden into the chondrules after the shock wave passed. Material not condensing into liquid droplets would later form the dust from which the matrix of the chondrite is formed. The chemical action of the primordial gas following the passage of the shock wave could have oxidized the iron in the fine dust particles, but not so in the larger—and thus protected—chondrules.

Some of the chondrules have suffered recrystallization. In these, as expected, some of the iron has become oxidized. It is supposed that chondrules which have been recrystallized became part of a larger body which was heated. This heating would account for the recrystallization.

It seems to be generally accepted that the round little chondrules embedded in the matrix of the chondrites were once molten. Although much study remains to be done on these very fascinating objects, their once molten condition, their chemical similarity to the sun, their free iron, make them unique in meteoritic studies. It is anticipated that they hold still more clues to the history of the solar system.

Whether chondrules are solidified droplets from the primordial gases of our solar system, and whether there is any connection between the minor planets and the meteorites is not known with certainty. But the entire story of the solar system cannot be learned in one day or one step. Speculation gives direction to observational

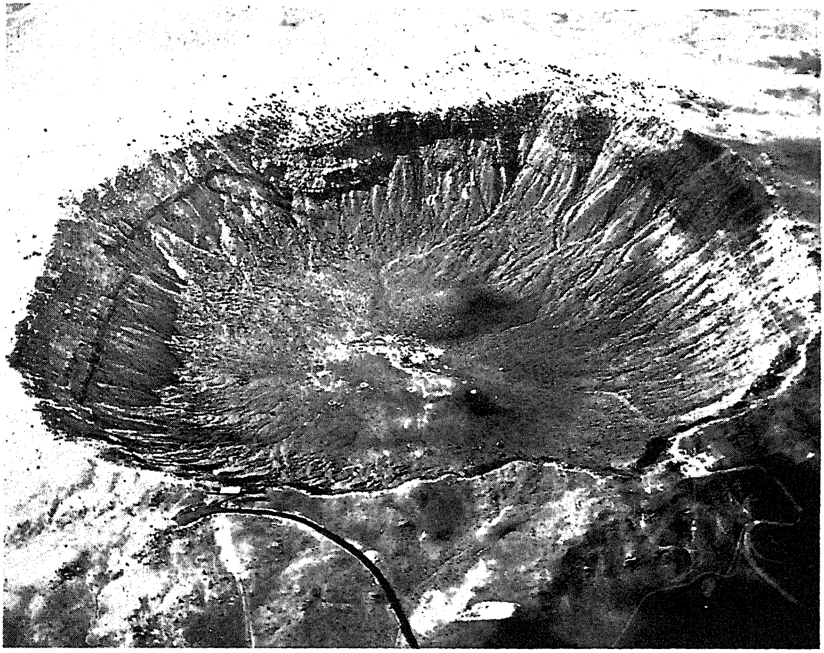


Figure 7-9 The Barringer Meteor Crater, Winslow, Arizona. (Meteor Crater Enterprises, Inc.)

research. If a speculation proves incorrect, progress has been made even though a new direction must be taken. Man has never seemed to be at a loss for new ideas and new directions, at least not in the past 400 years.

When a large meteorite falls, it leaves a crater on the surface of the Earth. Some of these craters are large, although none is as large as many of those on the moon. The largest that is definitely known to be of meteoritic origin is the Barringer crater near Winslow, Arizona (Figure 7-9), which is about 4,150 ft in diameter. The Chubb crater in Quebec, however, is nearly 2 miles in diameter and unlike any of the lakes in the surrounding area. These lakes were formed by glaciers and consequently are long and narrow. The lake in the crater, however, is quite circular. The crater is assumed to be meteoritic in origin, although, unlike the Barringer crater, no meteoritic fragments have been found around it. This may be because the area around the Chubb crater is covered with boulders which make it quite difficult to find fragments, whereas the area around the Barringer crater is flat and sandy.

The biggest fall in recorded history occurred in Siberia on June 30, 1908. The explosion resulting from this fall was big enough to make any manmade explosion seem small in comparison. At least ten craters were formed; presumably the meteor broke into smaller fragments when it hit the atmosphere. The trees outside the region of craters, to a distance of nearly 20 miles, were all blown down, with their tops pointing away from the explosion. Windows were broken for a distance of 50 miles, and an engineer about 450 miles away stopped his train in fright. The pressure wave traveled through the atmosphere as far as England, where it was recorded on barographs (instruments that measure atmospheric pressure). The shock waves traveled through the surface of the Earth and were recorded on seismographs (instruments that measure earthquakes) in Europe. Had the meteorite struck only 4 hours, 47 minutes later, the Earth would have turned enough so that it would have struck the city now called Leningrad!

Recent studies have shown that this explosion was too powerful to be accounted for by craters of the size found in the region of impact. It has therefore been speculated that it was caused by a comet and not by a meteorite. The gases in the comet, when heated in the Earth's atmosphere, could account for the magnitude of the explosion, for a gas such as methane would react chemically with the atmosphere.

That there are not more meteoritic craters on the Earth has caused some concern. Even with the geological changes, it seems that if the moon is so scarred and if Mars has been pelted by meteorites, then the Earth ought to have more such craters. Increased searching has revealed a number of *astroblemes*, fossil remnants of meteoritic craters. These scars of ancient collisions between the Earth and meteorites have been found on each of the great continents of the world. A number of them have been found in the Hudson Bay area of Canada in some of the oldest rock formations in the world.

On most nights during the year we see only *sporadic meteors*, but every now and then we are fortunate enough to witness a *meteor shower*. On these occasions there may be hundreds of meteors streaking across the sky in a single hour. Like eclipses, these showers frightened more superstitious peoples into thinking that the end of the world was imminent; but the world is still here with us, or we with it.

A photograph of meteors that constitute a shower makes it appear as though they originated in the same part of the sky like spokes from the hub of a wheel. This is the result of perspective. The point from which the trails radiate is called the *radiant*. The radiant effect can exist only if these meteors have been traveling in the same direction in space before they collide with the Earth's atmosphere.

To find the actual speed and direction of travel of meteors in space we observe their speed and direction of fall; from these we can subtract the velocity of the Earth and the effect of the Earth's gravitational field. It has been found that shower meteors travel around the sun in slightly elongated ellipses with periods of from 1.6 to 415 years for different groups. Such a group of particles is called a *stream*.

The Earth travels through one such stream about August 12 of each year, and since the radiant of this shower is in the constellation Perseus these meteors are called the Perseids. Their rate of fall averages about 50 meteors per hour. On October 9, 1933, there appeared a meteor shower (the Draconids) in which some 350 meteors were seen each minute at the peak of the shower. Again on October 9, 1946, the same meteor swarm was intercepted and another brilliant shower was observed. There are more than 25 meteor showers each year.

It is interesting to calculate how many meteors strike the Earth's atmosphere each day. This may be done by considering that any one observer will see only a small fraction (about 1/100,000) of the entire upper atmosphere. Since an observer sees about 10 meteors per hour on a night when no shower is in progress, the entire Earth is bombarded by about 24 million meteors each day. This figure is astonishing, to be sure, but even so it includes only those meteors that are visible. When we include the invisible ones the estimate runs to about 8 billion in one day. The number of invisible meteors can be roughly estimated by finding how many visible meteors of each size strike the atmosphere. The number increases regularly as the size decreases, and from this we can extrapolate to estimate the number of invisible meteors.

It might seem that so much meteoric material filtering through the atmosphere would add a considerable amount of mass to the

Earth. But each meteor is so small that the increase in mass is insignificant when compared with the Earth's mass. The increase becomes more understandable if expressed in weight; it has been estimated that as much as 5 million tons of extraterrestrial material reach the surface of the Earth each year. This is a sizable amount of material, but the calculated weight of the Earth is 7×10^{21} tons. At the present rate of collection, it would take roughly 10^{15} years to collect an amount of material equal to the mass of the Earth. Nevertheless, at this rate of collection, roughly 50 lb of material fall on each square mile of the Earth each year.

Comets

Only the comets have made a greater impression on mankind than the meteor showers and solar eclipses. To nearly every comet visible to the naked eye, in all but the most recent recorded history, man has been able to lay the blame for some disaster, whether it be war, massacres, the black plague, assassinations, floods, or earthquakes. Comets have been looked upon as omens of terror, blood, evil. Powerful men (who happened to live when a comet was seen in the sky) have been associated with them. The comet of 43 B.C. was supposed to have been the soul of Julius Caesar transported to heaven. This belief was encouraged by Augustus for political reasons. A comet was associated with the coming of William the Conqueror when it appeared in April of 1066. The comet of 1811–1812, of course, was called "Napoleon's comet."

It is true that comets are a spectacular sight when seen in the sky; little wonder that people who did not know what they were associated them with the evils that plagued man! When we learn more about comets, however, we find that they are of little danger to us; they do not strike the Earth very often, although one did strike it in 1908!

The first fruitful scientific work on a comet was done by Tycho Brahe, the astronomer who supplied Kepler with such a splendid list of planetary positions. Brahe compared the positions of the comet of 1577 as seen from Prague and from his island in the Baltic Sea, and by the method of parallax determined that it was farther away than the moon. Therefore it could not be an atmospheric phenomenon, as had been thought up to that time.

It remained for Halley, a contemporary of Newton, to predict the return in 1758 of a comet he had observed in 1682. Halley based his prediction on the orbit he had derived and on the fact that comets had appeared in the past at times that agreed with that orbit. He cited the comets of 1607, 1531, 1456, and 1305 as evidence that these were but passages of a single comet with a period of 75 or 76 years. On Christmas Day of 1758 Halley's comet did indeed appear in the skies, although he himself did not live to see his prediction confirmed.

A *periodic comet* revolves about the sun in an elliptical orbit which is usually quite eccentric. Halley's comet, for instance, comes within about 0.6 A.U. of the sun at perihelion and goes out as far as 35 A.U. at aphelion. Since all comets move in accordance with Kepler's three laws, they must travel faster at perihelion than at aphelion.* We are able to see comets with very eccentric orbits for only a small part of their journey because when they travel far from the sun they become too faint to be seen with our telescopes. Since a comet in a very eccentric orbit is observable only during a small part of its period, such an orbit is exceedingly difficult to compute. Our example, Halley's comet, has an orbit that brings it inside of Venus' orbit and takes it out beyond Neptune's.

Many of the periodic comets have orbits less eccentric than that of Halley's. Encke's comet, for example, has a perihelion distance of about 0.3 A.U., but an aphelion distance of only 4.1 A.U. Its period of 3.3 years is the shortest known. The orbit of comet Schwassmann-Wachmann lies between the orbits of Jupiter and Saturn and the comet is therefore observable every year.

Jupiter has such a strong gravitational influence on the comets that it has managed to "capture" a large number of them. There are about 60 comets whose periods are less than 100 years; these are termed the *short-period comets*. Some of them have aphelia that are close to Jupiter's orbit, comprising what is called "Jupiter's family of comets."

Although the planes of the orbits of all the planets are inclined by only small angles from one another, and their motions are all direct, the same is not true of comets, Halley's comet, for example, revolves in retrograde motion, and a number of comets have orbits that are nearly perpendicular to the ecliptic. The planes of near-para-

---

* As a rule of thumb, the velocity of a comet is comparable to the velocity of a planet at the same distance from the sun. Actually, since a comet's orbit is generally more elliptical than that of a planet, its velocity will be somewhat greater near perihelion and less near aphelion than that of a planet at the same distance.

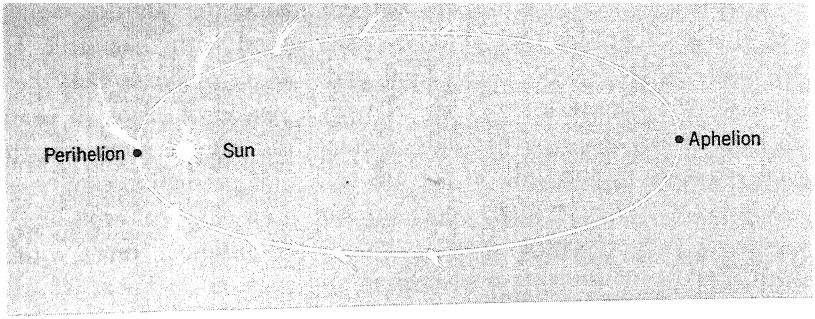


Figure 7-10 The comet's tail always points almost directly away from the sun.

bolic orbits are, more often than not, tilted by a large angle from the ecliptic, whereas the orbits of periodic comets conform more to the average plane of the solar system. The orbits of comets, then, are more erratic than those of minor planets, as those of minor planets are more erratic than the orbits of planets.

The aspects of comets are nearly as varied as their orbits. Among individual comets there are strong similarities and yet many variations. When viewed through a telescope, a comet will be seen to have a central *nucleus* which is quite bright and small. Surrounding the nucleus is a fainter and more nebulous structure called the *coma;* and this together with the nucleus composes the *head* of the comet.

Extending from the head is the *tail* of the comet, although unlike most tails in nature it is sometimes in front of the comet. The tail (and not all comets have tails) points away from the sun, so as the comet recedes from the sun the tail will lead the way. This seems contrary to what we might expect, but we must remember that in the space between the planets there is little that would cause the tail to trail behind as the resistance of the air in our atmosphere causes the smoke from a locomotive to trail.

As the comet approaches the sun the coma and tail grow larger. (Their growth is shown schematically in Figure 7-10 and pictorially in Figure 7-11.) The obvious cause of this effect is the sun.

The streams of particles (electrons, protons, and nuclei of other light atoms) comprising the solar wind are the cause of both the coma and the tail of a comet. To a lesser extent solar radiation makes a contribution. Gases are given off by the warming nucleus of the comet, are caused to emit their own light, and are driven away from

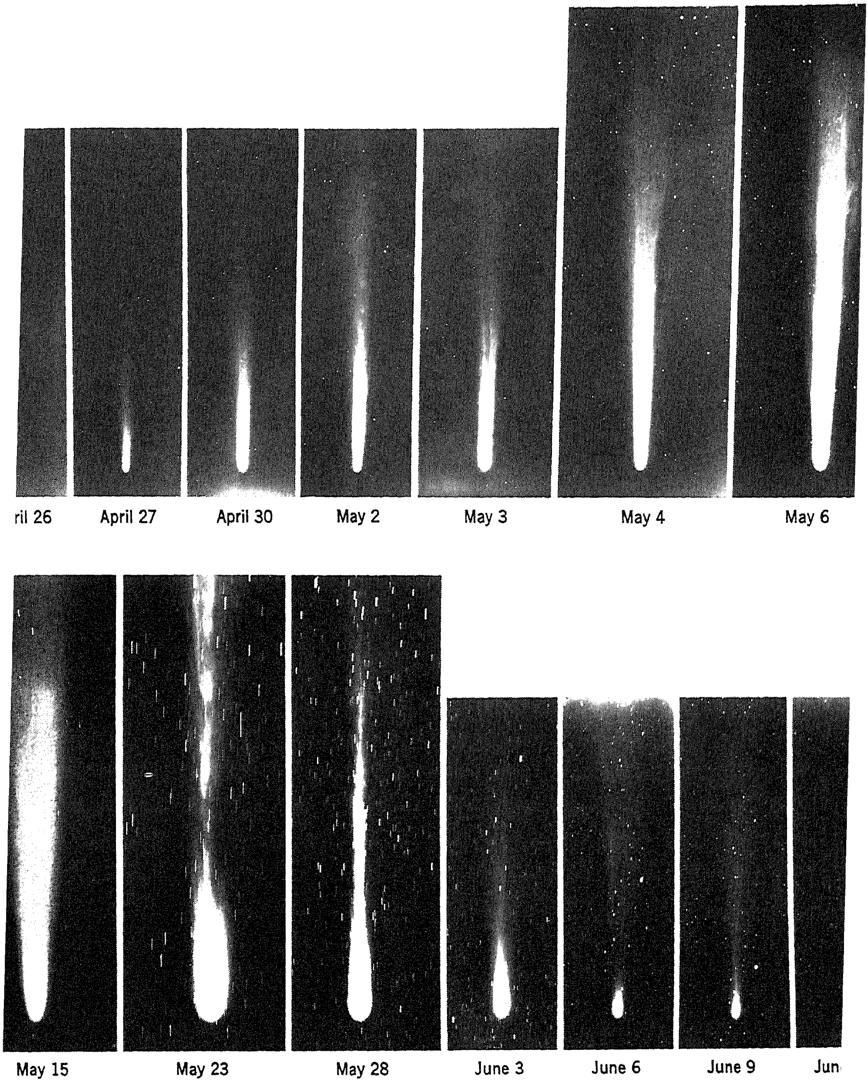


Figure 7-11 As Halley's comet approached the sun in 1910, its coma and tail grew in size. As the comet receded from the sun both the tail and the coma grew smaller. (Mount Wilson and Palomar Observatories)

the sun all by the solar wind. But studies of Comet Ikeya-Seki in 1965 indicate that only a thin outer shell of the comet is heated during a close approach to the sun. As the comet whirls about the sun, the gases in the tail are farther from the sun; they therefore lag behind, because each particle of gas assumes an orbit of its own.

Cometary tails will often show knots of material that progress outward from the comet, as with Comet Morehouse in October 1908 (Figure 7-12). The timing and velocity with which these knots move indicate that comets are subjected to sudden increases in the solar ejection of particles. These sudden increases in the solar wind are apparently caused by solar activity related to the sunspots and flares.

The size of the coma and tail is rather staggering. The tail of Halley's comet reached a maximum length of about 94 million miles, a little more than the distance of the Earth from the sun. The comet of 1843 had a tail that was nearly twice as long as this. The head of a comet (including the coma) may reach a diameter that rivals that of Jupiter; in fact, this is about the average size. The smallest coma observed was about 10,000 miles in diameter and the largest was that of Holmes's comet in 1892 with a diameter of 1,400,000 miles, nearly twice the diameter of the sun! However, the dimensions of any comet are not constant. The tail and coma are at their largest when the comet is near perihelion, but both shrink at great distances from the sun. They therefore do not give a complete picture of the comet, for both vary in size and composition. They do, however, help us determine the composition and structure of the nucleus.

If we analyze the light from a comet with a spectrograph we are able to learn much more about it. The nucleus gives a continuous spectrum which is apparently reflected sunlight. The coma and tail yield an emission spectrum composed of both the bright lines caused by atoms and bright *bands* from molecules. (Molecules, because they are several atoms together, emit a great many lines in groups which are called bands.) These molecules absorb some of the energy of the sunlight and reradiate it. Many of these molecules detected in comets, however, are not the ones we are likely to find on the Earth or any other planet. Instead, we find such strange molecules as C_2 (carbon), CN (cyanogen), NH (nitrogen hydride), CH (methylidyne), and OH (hydroxyl). In addition, there are a few common molecules, such as CO (carbon monoxide) and N_2 (nitrogen). Some of these have become ionized, that is, they have lost one electron (for example, CH^+, CO^+, OH^+, and N_2^+).

The strange molecules are incomplete and are called free radicals. If any one of them were placed in the Earth's atmosphere it would quickly grab one or more atoms and by doing so become a stable molecule. Under very low pressure, however, atoms are too far apart for an incomplete molecule to capture free atoms. The fact that these

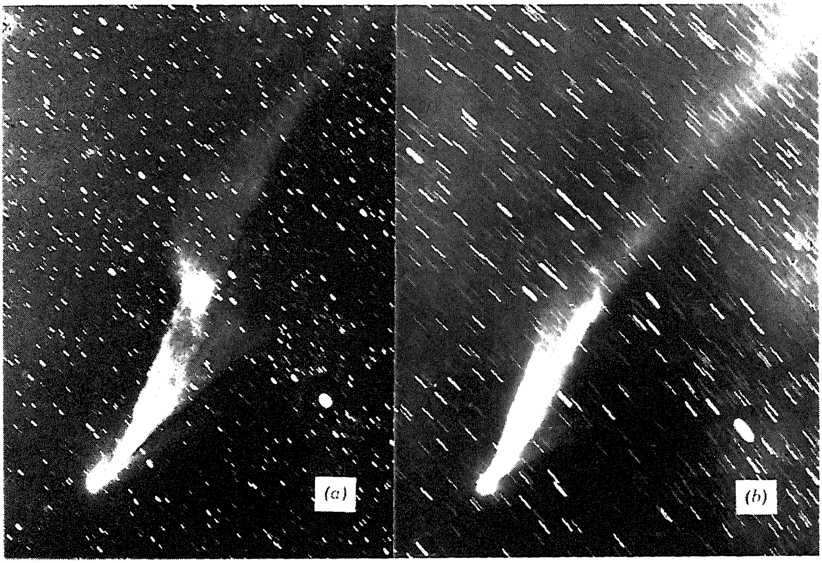


Figure 7-12 Comet Morehouse showing changes within 3 hours on September 30, 1908. (Yerkes Observatory)

strange molecules are found in the gases composing the coma and tail of a comet, then, leads to the conclusion that these gases are in what would be considered a vacuum by terrestrial standards.

Since the gases that leave the comet by way of the tail never return, and since their supply is not inexhaustible, the comet becomes less spectacular after repeated trips near the sun. The bright "naked eye" comets that appear unexpectedly from time to time have apparently made only a few trips around the sun, whereas the short-period comets that spend more of their time close to the sun have very short tails or none at all.

It is of interest to consider the nucleus of the comet, since both coma and tail originate in this small and less conspicuous part. From the nucleus we obtain a spectrum that is reflected sunlight, and we therefore assume that it is denser than the gaseous portions. Beyond this we are left with little to go on. It was thought that when Halley's comet passed directly between the Earth and the sun in 1910 its silhouette would be visible as it transited the sun. But nothing was seen!

The comet's tail was to sweep past the Earth; this, it was hoped, would enable astronomers to make observations of any reaction between the tail and the Earth's upper atmosphere. However, the Earth's passage through the tail occurred on May 21, 1910, when the moon was unfortunately nearly full; if anything unusual happened it was washed out by the bright moonlight.

It had been announced to the newspapers that the tail consisted of poisonous gases, an announcement, which when interpreted by the newspapers and the public, led to a minor form of mass hysteria and fright for fear that the world would come to an end. It not only initiated the sale of nostrums reputed to confer immunity from the noxious effects of the comet's tail but is thought to have been the cause of a number of suicides. But the tail of the comet, which terrified people then as in centuries past, could be neither seen nor detected when the Earth was quite in the midst of it.

Most of a comet's mass obviously resides in the nucleus, but this mass is very difficult to determine. The best indication of a comet's mass was obtained by Dr. Elizabeth Roemer at the University of Arizona. It was noticed in 1957 that Comet Wirtanen, 1956*c* (the *c* indicates that it was the third comet discovered in 1956) had split into two parts (Figure 7-13). By continued observation over 2 years, Dr. Roemer was able to determine the velocity with which the two nuclei separated. Making the assumption that this velocity was the velocity of escape of the nuclei, she was able to estimate that the total mass of the comet was about 10^{14} kilograms. On the Earth this would weigh about 10^{15} pounds. A rock with that weight on Earth would have a diameter of about 2 miles. A ball of ice and snow would have a diameter closer to 3½ miles.

We do have one clue to the structure of a comet's nucleus. In a preceding section it was mentioned that meteors sometimes appear in showers, and that these showers result from myriads of tiny meteorites called *micrometeorites* traveling in a common orbit about the sun. When a careful study is made of these orbits it is found that nearly all agree quite closely with the orbit of a particular comet. For example, the shower that occurs about August 12, the Perseids, follows the orbit of comet 1862 III; the Leonids, which form a shower about November 14–18, have the same orbit as Comet 1866 I; the May Aquarids are associated with Halley's comet; and the Draconids of October 9 are closely associated with the comet Giacobini-Zinner. In fact, on October 9, 1946, the Earth missed the latter comet by only 140,000 miles; that is, we passed through its orbit only 15 days after the comet passed the same point in space.

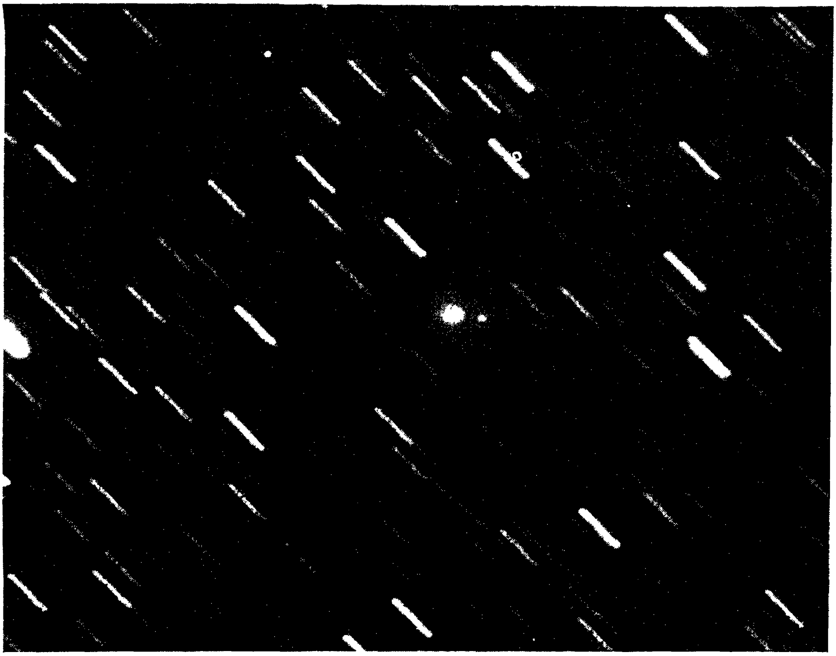


Figure 7-13 Comet Wirtanen, 1956 *c*, showed a double nucleus in 1957 and 1958. This photograph was taken on April 27, 1958, at the U.S. Naval Observatory. Flagstaff, Arizona. (Official U.S. Navy Photograph)

It therefore appears fairly obvious that meteor showers result from the Earth's attracting and sweeping up debris that was left in the orbit by a comet. The comet, then, must be composed of these tiny particles and the nucleus could be a swarm of micrometeorites trailing its stragglers like a swarm of bees.

But the gases that form the coma and tail could not originate from micrometeorites; the comet is too cold to vaporize them. It is therefore postulated that the micrometeorites of the comet's nucleus are held together by a spongy mixture of ice, frozen methane, ammonia, and perhaps carbon dioxide and C_2N_2. As the comet nears the sun some of this frozen mixture begins to vaporize, yielding the coma and the tail. The coma is composed of the gases that have evaporated from the nucleus, and the tail is caused by the action on these gases of the pressure of radiation and atomic particles from the sun.

Hundreds of meteorites have been collected, but until recently no micrometeorites have been collected; they are too small to burn in the upper atmosphere. They are simply slowed by the air and drift down as undistinguished dust. A group of astronomers headed by C. L. Hemenway at the Dudley Observatory in Albany, New York, have collected micrometeorites by the use of rocket and balloon flights. Collecting panels of nitrocellulose or of aluminum foil coated with mineral oil were sealed in dust-free boxes. Every effort was taken to make sure that the panels contained only material collected above the atmosphere: 100 miles for the rockets and 20 miles for the balloons. At the proper height the sealed boxes were first opened and then closed by a radio signal. The micrometeorites collected are very small, on the order of 10^{-3} cm. Some of them are strange jagged shapes, others are spherical, still others are fluffy. Balloon flights made during a meteor shower have collected up to 80 times the number of micrometeorites collected when no shower was in progress. There seems to be every reason to believe that these particles were once part of a comet!

It is becoming clear that the sporadic meteors and the meteorites should be distinguished from the meteor showers and the micrometeorites. The sporadic meteors and meteorites apparently come from the minor planets; some of them may result from the catastrophic collision of two larger ones, but some may result from collisions of lesser minor planets. The micrometeorites, on the other hand, are apparently cometary debris and have nothing to do with the minor planets.

Famous Comets

A discussion of some of the more notable comets will help to emphasize and clarify some of the features already mentioned. The most famous, Halley's comet, has been mentioned already, but there are a few interesting facts concerning it that remain to be told.

Since Halley's comet is so bright, it has been seen by many peoples in recorded history. Only one passage since 240 B.C. has not been found in historical or archeological records, many of which were made by the ancient Chinese astronomers. It has made twenty-nine passages close to the sun and it will again appear as a spectacular sight in the sky in 1985 or 1986. There are indications that at one time it was more spectacular than in its more recent visits. But it is dangerous to take literally the descriptions of comets given by ancient peoples; they were too awed and frightened by the sight to speak objectively of the appearance.

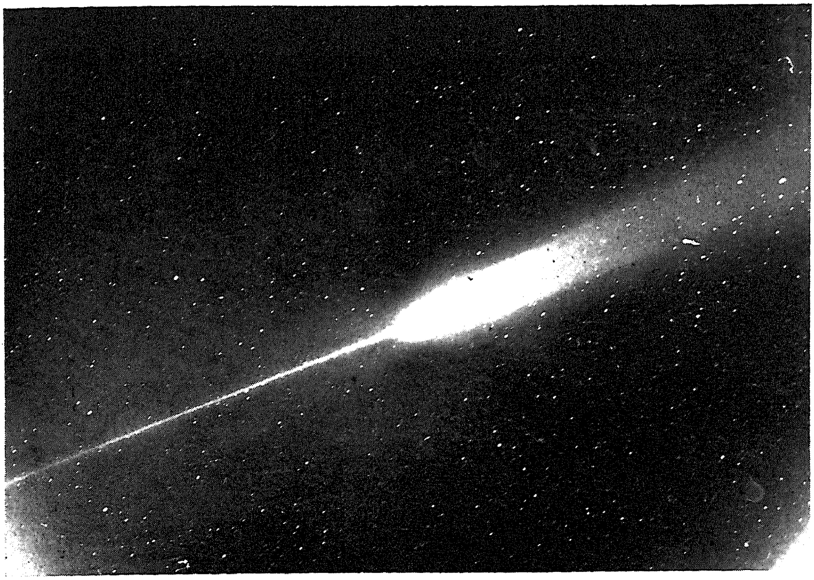


Figure 7-14 Comet Arend-Roland showed a long "spike" pointing towards the sun when the Earth passed through the plane of its orbit on April 25, 1957. (Uppsala Observatory)

Two unexpected comets which appeared in 1957 became bright enough to be seen by the naked eye and to make the year an unusual one. The first was Comet Arend-Roland (1956*h*) which was discovered on November 18, 1956, at the Uccle Observatory, Belgium, by Sylvain Arend and Georges Roland. It was followed until February when it went behind the sun. Observations made it possible to predict that when the comet came out from behind the sun it would be bright enough to be seen without the help of the telescope. By April 23, 1957, many people had seen it in the northwestern sky shortly after sunset, and it remained visible for some weeks although it grew fainter and fainter as it receded from the sun and the Earth.

This comet was unusual not only because it was so bright but also because of its "spike," which appeared to be an "antitail" pointing toward the sun. There is no known reason why a jet of gas should point toward the sun, since particle and radiation pressure will push all the gases away from the sun, and gravity does not play favorites. The "antitail" looked somewhat like a fan when first seen on April 22–23, but on April 25 (Figure 7-14) it appeared as the "spike,"

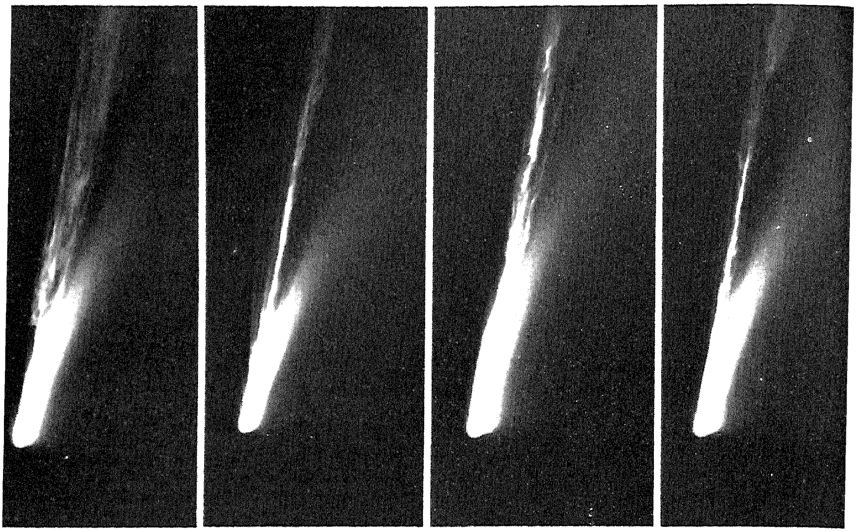


Figure 7-15 Four views of Comet Mrkos taken with the 48-in. Schmidt telescope on Mount Palomar. (Mount Wilson and Palomar Observatories)

long and narrow. This feature is thought not to be actually a spike pointing toward the sun but merely sunlight reflected by the debris that was left behind in the comet's orbit. According to this view, when the Earth passed through the plane of the comet's orbit on April 25, the debris was seen edgewise; it thus appeared brighter and narrower than at other times.

We might think that one naked-eye comet would be enough for one year, but no; shortly after Comet Arend-Roland faded from view there appeared Comet Mrkos, technically known as 1957*d*. Apparently first seen on July 29 by Sukehiro Kuragano in Japan, it was again noted on July 31 by an airline pilot, Peter Cherbak, as he was flying over Nebraska. Antonin Mrkos of Czechoslovakia sighted the comet on August 2; since he was the first to report it officially it was given his name. Comet Mrkos had an unusual tail in that it was broken into two main parts: one part seemed to be straight but to exhibit occasional kinks as though it were wagging; the other arched away from the straighter tail (Figure 7-15). The formation of more than

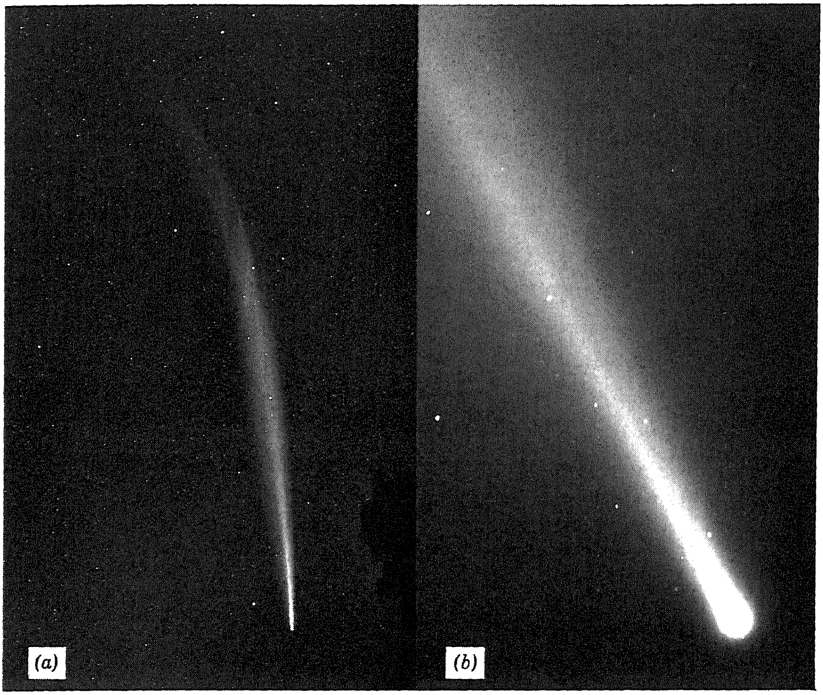


Figure 7-16 (*a*) Comet Ikeya-Seki as photographed with 35 mm camera. (Courtesy Elizabeth Roemer). (*b*) The head of comet Ikeya-Seki taken with the 40-in. reflector at the U.S. Naval Observatory, Flagstaff, Arizona at nearly the same time that (*a*) was taken. (U.S. Navy)

one tail, as in Comet Mrkos, is apparently caused by the fact that molecules may or may not be ionized. Molecules with no electric charge will move differently through the magnetic field of the solar system than molecules with an electric charge.

The fabulous Great Comet of 1965, Comet Ikeya-Seki, was discovered on September 18, 1965, by two Japanese amateurs K. Ikeya and T. Seki (Figure 7-16). This exceptional comet had a perihelion distance of only 290,000 miles from the surface of the sun and a perihelion velocity of 300 miles per second! That the comet survived such a close approach to the sun (recall that the moon is 240,000 miles from the Earth) indicates that really splendid insulation properties operate near the surface of the comet. It was expected that the tidal forces from the sun would affect the comet in some manner; on November 4, 1965, just 2 weeks after perihelion passage, its nucleus was observed to be split into two pieces. There were even reports of a third nucleus. The fact that cometary nuclei split into smaller parts is direct evidence that they are not solid in the same sense that a minor planet is.

Interplanetary Dust

The material that leaves comets may in part account for the myriads of tiny particles that pervade the solar system. These particles become visible just after sunset or before sunrise as a faint glow extending up into the sky from a point on the horizon just above the sun. The faint glow follows the ecliptic and is called the *zodiacal light,* for it is seen along the zodiac (the band of twelve constellations along the ecliptic). The ecliptic is the plane of the Earth's orbit and is also nearly coincident to the plane of the entire solar system; consequently as we look along the ecliptic we see more of the particles reflecting the sun's light. (It thus resembles the spike of Comet Arend-Roland.)

Observations tell us that the zodiacal light is not redder than the sun's light, which indicates that the particles reflecting the sun's light are larger than minute dust particles. Perhaps we see the micrometeorites reflecting the sunlight.

The zodiacal light is partially polarized; this indicates that electrons reflecting sunlight contribute some of the total amount of light seen. There is every reason to believe that the zodiacal light is part of the sun's vastly extended atmosphere. The sun's outer atmosphere is called the corona and it is best seen during a total solar eclipse. The brightness of the corona decreases with increasing distance from the sun, and this can be expressed mathematically. The brightness of the zodiacal light also decreases with increasing distance from the sun, and by the same mathematical relationship used to describe the corona!

Around midnight there can be seen in the sky directly opposite the sun, a glow similar to, though fainter than, the zodiacal light; it is called *Gegenschein* (German, counterglow). This is suspected of being the sun's light reflecting from micrometeorites too.

In order to see either the zodiacal light or *Gegenschein* the observer must look into a sky that does not receive light from cities, towns, or even the moon. Zodiacal light is fainter than the milky Way, and *Gegenschein* is so faint that some people cannot see it under the best of conditions.

We see that the entire composition of the solar system includes many different types of bodies: the sun, a hot gaseous sphere much larger than any planet; the Jovian planets (Jupiter, Saturn, Uranus, and Neptune), composed principally of hydrogen and helium, the second largest bodies in the system; the terrestrial planets (Mercury, Venus, Earth, and Mars), which are rocky, metallic, and consequently denser than the Jovian planets; the satellites that revolve about the planets; the minor planets, which lie mainly between the planets Jupiter and Mars; the comets which may be frozen gases impregnated with small particles; and the tiny particles and gases that pervade the entire solar system.

Basic Vocabulary for Subsequent Reading

Achondrite
Chondrite
Chondrule
Comet
Fireball
Gegenschein
Meteor
Meteorite
Micrometeorite
Minor planet
Radiation pressure
Tektites
Zodiacal light

For Further Reading

Hawkings, Gerald S., *Meteors, Comets and Meteorites,* McGraw-Hill Book Co., New York, 1964.
Mason, Brian, *Meteorites,* John Wiley and Sons, New York, 1962.
Nininger, Harvey H., *Out of the Sky,* University of Denver Press, Denver. Colo., 1952.
Richter, Nikolaus B., *The Nature of Comets,* Dover Publications, New York, 1963.
Anders, Edward, "Diamonds in Meteorites," *Scientific American,* p. 26, October 1965.
Barringer, D. M., "The Meteorite Search," *Natural History,* p. 56, May 1964.
Beals, C. S., "Fossil Meteorite Craters," *Scientific American,* p. 32, July 1958.
Dietz, Robert S., "Astroblemes," *Scientific American,* p. 50, August 1961.
Florensky, Kirill P., "Did a Comet Collide with the Earth in 1908?" *Sky and Telescope,* p. 268, November 1963.
O'Keefe, John A., "Tektites and the Cyrillid Shower," *Sky and Telescope,* p. 4, January 1961.

O'Keefe, John A., "Tektites and Impact Fragments from the Moon," *Scientific American*, p. 50, February 1964.

Livingston, et al., "The Sodium D Lines of Comet Ikeya-Seki," *Sky and Telescope*, p. 24, January 1966.

Marsden, Brian G., "The Great Comet of 1965," *Sky and Telescope*, p. 332, December 1965.

Mason, Brian, "The Unearthly Visitors," *Natural History*, p. 18, April 1961.

Pettersson, Hans, "Cosmic Spherules and Meteoritic Dust," *Scientific American*, p. 123, February 1960.

Wood, J. A., "Chondrites and Chondrules," *Scientific American*, p. 65, October 1963.

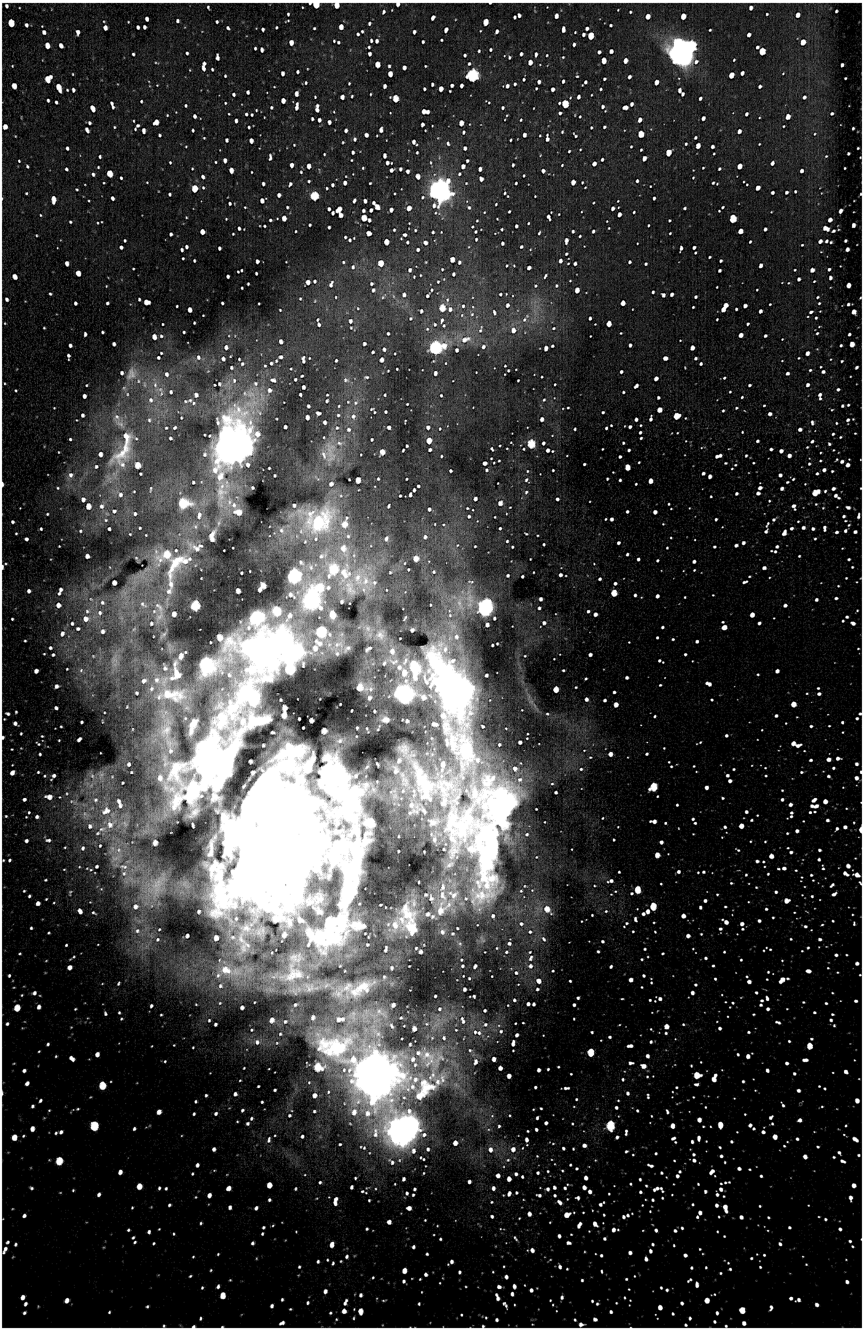
"Bosomtwe: An African Meteorite Crater?" *Sky and Telescope*, p. 15, July 1965.

"Micrometeorite Collecting with High Altitude Rockets," *Sky and Telescope*, p. 84, February 1962.

"Origin of Some Iron Meteorites," *Sky and Telescope*, p. 282, May 1965.

"Shower Meteorites Collected by Balloon," *Sky and Telescope*, p. 276, November 1965.

"Two More Ancient Canadian Meteorite Craters," *Sky and Telescope*, p. 198, October 1963.



The Lagoon nebula (M8) with dark "globules" of gas and dust that could perhaps form new stars with planets about them. (Lick observatory)

Chapter 8 The Age and Origin of the Solar System

Man's speculations and researches into his origins have led him to theories of evolution, both for himself and for the entire universe. The concept of evolution implies a beginning and leads to such questions as: "When and how did the Earth come into being? From where came the moon and the sun, and the comets, meteors, stars?" Peoples of all ages and races have asked these questions and have answered them to the best of their abilities and according to their beliefs. But none of the answers given in past ages satisfy the requirements of modern scientific thought and observation. Neither, for that matter, has any claim to complete accuracy yet been made for our modern description of the beginnings and evolution of the universe as a whole or of the solar system in particular. But the search alone has its fascinations, stimulated as it is by the hope that ultimately the story of the universe may unfold from a mass of data and a few germinative ideas.

Age of the Solar System

How old is the solar system? The question is basic to any theory of origins, but how is it to be answered? Age reveals itself in many different ways. The age of a tree is determined by the number of rings in its trunk, that of a horse by its teeth. If there is some analogous sign or evidence in the solar system that will help us determine its age we may expect to find it first at home, on the Earth.

To obtain evidence of the Earth's age we must first find some material or object that shows signs of change over a long period of time. Then to determine this period, three factors must be considered:

1. The material's present condition.
2. Its assumed condition when formed.
3. Its rate of change.

Rain water, for instance, is nearly pure (that is, fresh); the oceans are salty; and the Great Salt Lake and Dead Sea are even saltier. The rain falls on the land and carries salts from it into the ocean. The ocean's water evaporates, leaving the salt behind, to form clouds that cause rain to wash again down the mountain slopes and deposit more salt into the oceans. The cycle goes on; how long has it been doing so?

The ocean's present condition is one of water with dissolved salts. It is reasonable to assume that its condition when formed was one of fresh water, as fresh as river water. Its rate of change from a fresh to a saline condition may be deduced from the amount of salt carried into the oceans each year by the present rivers. If we divide the number of tons of salt in the oceans at the present time by the number of tons deposited each year, we obtain the length of time this process has been going on. From this calculation it has been estimated that the oceans are about 300 million years old.

But the geologists warn us to be careful because the Earth that we now see is *not* typical of its past history. There are many more rugged mountains now than in former geological epochs. This means that the rivers erode the Earth's surface more rapidly and thus carry more salt than usual into the oceans. Consequently, we postulated too rapid a rate of salt cumulation and 300 million years turns out to be too short a time interval. When we take into consideration the slower rates of cumulation before the present era, we arrive at an age for the oceans that is closer to a few billion years.

This method, approximate though it may be, does enable us to be certain of one thing: the oceans are not infinitely old; they did begin sometime. If salt had been deposited in them for an infinite length of time they would have become completely saturated and could have held no more salt, for they are finite in size and therefore the amount of salt required to saturate them is also finite. If the oceans had become saturated with salt any additional salt would have settled to the bottom. If this had happened we would have to investigate the amount of sediment on the ocean floors to arrive at an estimate of the length of time that sedimentation has been going on.

The age of the oceans, however, need not be the same as the age of the Earth. Indeed, we expect that they would be younger, for the assumption has been made (and not without good reason) that the surface of the Earth was originally molten and much too hot for any water to collect or even for any rain to fall. We must look for some other change that will give us a clue to the age of the Earth itself.

One branch of physics (nuclear physics) concerns itself partly with radioactivity. A radioactive element is one whose atoms are unstable and therefore break up or *decay* into other atoms. It has been found that the rate of radioactive decay is constant and quite independent of any pressure or temperature that man has imposed on any radioactive substance. Some of these elements appear naturally in the rocks of the Earth, and the fact that they decay at a known rate gives us a clue to the age of the rocks.

Uranium 238 (the 238 is the atomic mass of this particular kind of uranium) is one of these natural radioactive elements. It decays into another radioactive element, thorium 234, which decays into yet another radioactive element, protactinium 234, and so on through a number of radioactive elements until lead 206 is reached. Another naturally occurring radioactive element, thorium 232, decays into lead 208, another stable isotope of lead. Neither lead 206 nor lead 208 decays into any other element. Furthermore, lead is a solid, and consequently accumulates in the rock that contains the uranium 238, just as salt accumulates in the oceans.

The rate of decay of a radioactive element is generally measured by the *half-life* of that element. The half-life of any radioactive element is the time it takes for half of the original atoms to decay.

For example, the half-life of the uranium 238–lead 206 process is 4.5 billion years, which means that no matter how many atoms of uranium 238 we start with, in 4.5 billion years half as many atoms of lead 206 will have been formed. The half-life of the thorium 232 to lead 208 series is 1.3×10^{10} years.

To find the age of a particular rock we first determine its present condition by measuring the amount of lead and uranium in it. We can simplify this example by assuming that the rock contained no lead when it was formed; therefore all the lead it now contains comes from the radioactive decay of uranium. We can then obtain the age of the rock, that is, the length of time since it solidified, by applying the known rate at which lead has been formed through the decay of uranium.

This sounds very easy. Unfortunately, however, we were not justified in assuming that the rock started without lead; if there was uranium in the primordial rock there may certainly have been some lead. In fact, we are confronted with the initial problem of determining the amount of primordial lead.

Harrison Brown, of the California Institute of Technology, has attacked this problem by recognizing and investigating the fact that certain sections of meteorites contain no radioactive elements. By measuring the amount of lead in these sections he deduced the amount of primordial lead in the meteorites studied. From this and other measurements he estimated the age of the *meteorites* he studied to be 4.6 billion years. Other scientists have used different methods and determined the age of some terrestrial rocks to be 3.4 billion years, although not all rocks are of the same age. Since the terrestrial rocks that have been studied have different ages we cannot be sure that the oldest one has yet been found. On the other hand, all the meteorites studied are of more nearly the same age, as well as being older than any rock yet found on the Earth. The age of the solar system is therefore generally taken as 4.6 billion years.

The age of the solar system, important though it may be, is only part of the evolutionary story; the question "how?" still faces us. Here we shall discuss only the most important of many theories of the solar system's origin.

Origin of the Solar System

A complete description of the origin of the solar system must be able to explain the main features of the system as it now exists, unless we assume (though it is very unlikely) that there has been some strong outside disturbing force since the origin. In general we could say that the solar system is fairly flat with almost everything spinning in the same direction. The following are the specific features which such a description must explain:

1. The motions (revolution and rotation) of nearly all the bodies that revolve about the sun are in the same direction, except those of Uranus and Venus, some of the satellites, and some of the comets.
2. The planes of revolution of these bodies are within a small angle of one another. The exceptions are the satellites of Uranus and some of the comets.
3. The planes of rotation of the planets as well as that of the sun are nearly the same as the planes of revolution of the planets. Uranus is again an exception.
4. The densities of the four terrestrial planets are considerably greater than those of the four Jovian planets (Table 8-1).

One of the early hypotheses (that of the Frenchman Buffon) was based on the belief that a comet is a huge mass of material that, if it were to strike the sun, would surely cause debris to fly out into space. The Earth and planets would be formed of such debris.

Table 8–1 Densities of the Planets in Terms of Water

Terrestrial Planets		*Jovian Planets*	
Mercury	5.5	Jupiter	1.3
Venus	4.9	Saturn	0.7
Earth	5.5	Uranus	1.6
Mars	3.9	Neptune	2.3

We now know that a comet's mass is too small for this to be possible. The hypothesis, however, was revised; instead of colliding with a comet, the sun might have collided with another star. To be sure, if two stars were to collide, things would happen on a grand scale. But the stars are immensely distant from each other, and the possibility that a collision might occur between them is so remote that the idea may be dismissed on these grounds alone.

If a collision is out of the question, what about the close approach of another star? Two stars might come so close to one another that tidal actions set up in each star might draw out a cigar-shaped blob of gas from the sun (or both stars) which might eventually condense into the planets and satellites. But the probability of this happening is also extremely remote; it is estimated that during the lifetime of the Earth only ten planetary systems could have formed by this process in our Milky Way galaxy of 100 billion stars.

The astronomer, faced by billions of stars, cannot fall heir to the belief that the Earth is something special, the planetary system unique. If the universe is to be investigated objectively at all, he cannot accept the belief that all the billions of stars exist solely for the benefit of mankind. He assumes that there are many other stars circled by planets that support life not too unlike our own. And for this reason he is very critical of hypotheses, like those of stellar collision or close approach, whose acceptance would imply the uniqueness or near-uniqueness of our solar system.

Furthermore it has been shown that even if two stars were to collide or suffer a close approach, the gases spewed out into space would not condense into planets; the density would be too low and the motions would only lead to further dispersal of the gas. Therefore, astronomers have sought to explain the formation of the solar system as a self-contained event with no outside influences.

Nebular Hypothesis

The hypotheses which satisfy the restrictions of self-containment have one basic similarity: they all base their arguments on the supposition that the sun originated in a great nebulous cloud of gas and dust. *Gaseous nebulae,* composed primarily of hydrogen gas, are very prevalent in our Milky Way galaxy. It is believed possible that a portion of such a nebula might become localized to the extent that it separates from the main body of gas by becoming a local "whirlpool" or *eddy.* Photographs of some bright nebulae that reveal dark "globules" of gas between them and the Earth add weight to this argument (see the frontispiece of this chapter).

Now, if the gas in such an eddy has a density that is high enough it will be able to contract under its own gravitational influence. Any object or system that is rotating *and* at the same time contracting will have to rotate more rapidly still as a consequence of the law of *conservation of angular momentum.*

Angular momentum is a difficult quantity to imagine: for a given particle in a gaseous eddy it is equal to the product of the particle's mass, the radius of the orbit in which the particle moves, and the velocity of the particle ($AM = mrv$). If no outside forces act on the particle, this product must remain constant.

In this case the particle's mass does not change. Therefore, since the radius of its orbit decreases (because the entire eddy is contracting) its velocity must increase. The phenomenon may be illustrated by the example of a spinning ice skater. As she starts to spin her arms are outstretched; then as she brings her arms in close to her body she spins more rapidly.

If the radius of the orbit decreased and the velocity did not increase, the angular momentum would also have to decrease. But the law of conservation of angular momentum denies this possibility—if no outside forces act on a rotating system, its angular momentum must remain constant.

As a result of the more rapid rotation of a contracting eddy in a nebula, then, a large disk of gas would form around a central nucleus in a plane perpendicular to the axis of rotation. The result would appear much like a slightly out-of-focus picture of Saturn and its rings. At this stage the sun is a new star which has a very large and flat disk of gas and dust around its midriff. The flatness of this disk would account for the fact that the planetary orbits lie nearly in the same plane.

According to Kant-Laplace. From this point on, various hypotheses concerning the origin of the solar system begin to diverge. The first to postulate a gaseous disk were the German philosopher Kant and the French mathematician Laplace. Once the disk had formed, Laplace suggested, it would in turn contract. In so doing it would periodically leave smaller rings or "zones of vapors" which would become detached from the main disk by ceasing to contract. Each of these detached rings would then coalesce into a planet by forming a little eddy of its own; and the rotation of this eddy would in turn form smaller gaseous rings from which the new planet's satellites would form. Laplace cited Saturn's rings as "existing proofs of the primitive extension of the atmosphere of Saturn, and its successive condensations." Ingenious as this explanation is (it partially anticipated contemporary hypotheses), it fails to account for one important aspect of the solar system: the very unequal distribution of angular momentum between the planets and the sun. Although the sun is much more massive than the rest of the solar system (it contains 99.9% of the total mass), it comprises only 2% of the angular momentum! Physically this means that the sun rotates quite slowly while the tiny planets with 98% of the system's angular momentum revolve at great speeds and great distances from the center.

The validity of the theory of Kant and Laplace can be tested by using it, together with additional knowledge gained since their time, to predict a period of rotation for the sun. Not only have astronomers been able to estimate the size of the gaseous cloud before it contracted, but they have also been able to measure the speed of rotation of gases in observed nebulae. Using this assumed size, the observed speed of rotation, and the law of conservation of angular momentum, the Kant-Laplace theory predicts that the period of rotation of the sun should be about ½ day; the observed period of rotation is about 26 days. Such a large discrepancy between theory and observation is not acceptable; the Kant-Laplace theory must omit some important aspect of the process of formation. Even though it has proved a fertile guide to the thinking of later astronomers, it must be discarded.

According to von Weizsäcker. The first contemporary hypothesis to be based on nebular contraction, with a disk about the sun, was proposed in 1944 by the German astronomer C. F. von Weizsäcker. Through a mathematical analysis he established that in any gaseous disk surrounding the sun the particles would move, *with respect to each other,* in orbits that are shaped like kidney beans. In accounting for the observed distances of the planets from the sun, von Weizsäcker assumed that five "kidney bean" orbits or eddies would form in strings around the sun at specified distances (Figure 8-1). Each of these strings of five eddies would in turn revolve about the sun because the original nebula rotated. Thus each particle in a given eddy would actually revolve, with respect to the sun, in an elliptical orbit whose perihelion distance would be equal to the distance from the sun to the inside of the eddy (p) and whose aphelion distance would equal the distance from the sun to the outside of the eddy (a).

Weizsäcker postulated further that the planets were formed in the gaseous disk by *accretion,* the process by which a larger body grows by collecting many smaller particles. Since accretion must begin when the particles are so small that their mutual gravitational pull would be ineffective, the theory demands that the particles not only collide but also stick together. Collisions of particles are most likely to occur in the region between two adjacent strings of eddies, for here particles travel in opposite directions. This can be vouched for by any person who has inadvertently turned into a one-way street going in the wrong direction.

A collision between two particles of about the same size would be apt to shatter them both. But if a smaller particle were to strike a larger body it might well stick to it, contributing to its growth. After a collection of such particles became larger its gravitational field would be stronger and more able to attract other particles. Hence the process of accretion would proceed more and more rapidly until

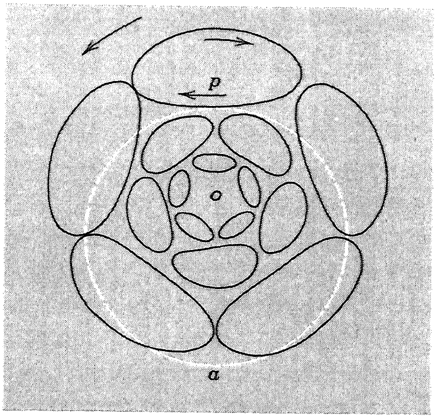


Figure 8-1 The motions of particles in the original solar nebular disk depend on the frame of reference. If the motions are compared to the sun, any given particle would revolve in an ellipse (dashed line); *p* is perihelion and *a* is aphelion. But when their motions are compared to each other, any given particle would revolve in an ellipse (dashed line); *p* is perihelion and *a* sun. The arrows indicate the direction of rotation of the disk itself (the outside arrow) and of the eddies (the inside arrows). Von Weizsäcker considered strings of five kidney bean orbits about the sun.

the disk's supply of particles had diminished.

This process favors the growth of a few large bodies, the planets, between two strings of eddies. Each eddy rotates in one direction while revolving about the sun in the opposite direction. The growing planet, called a *protoplanet,* is caught between two strings of eddies and reacts like a roller bearing, rotating in the same direction in which it revolves. Thus the direction of rotation of the planets is the same as the direction of their revolution.

According to Kuiper. A second hypothesis based on the idea of a nebular disk about the newly formed sun has been proposed by G. P. Kuiper, of the Steward Observatory of the University of Arizona, who argues that the "kidney bean" eddies would be more irregular, and that the larger eddies would last for a longer period of time. It is even possible that smaller eddies might form inside the larger ones (Figure 8-2).

This change from a regular pattern to an irregular one disposes of one objection to von Weizsäcker's description, namely, that dust particles will not stick together; if they did, how could we have dust storms?

Kuiper bases his idea of a planet's beginnings not on accretion but on the fact that two opposing forces act on an eddy of gas in

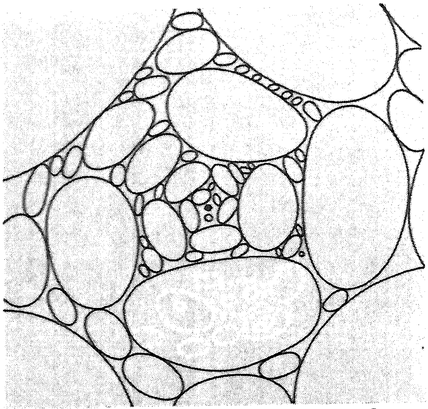


Figure 8-2 Kuiper suggested that the kidney bean orbits about the sun may not have been regular and in fact may have even changed, the big ones giving rise to smaller ones, which then grew in size. (By permission from *Astrophysics*, edited by Hynek, chapter by Kuiper. Copyright 1951, McGraw-Hill Book Co.)

such a nebular disk: the force of its own gravity, which tends to pull it together, and the gravity of the sun, which sets up tidal forces that tend to pull it apart. If the density of the gas is high enough (higher than a critical value Kuiper called the Roche density), its gravitational field will gain control, and the material composing the eddy will contract into a protoplanet with a central nucleus surrounded by a disk of gas much as Laplace envisioned.

The tidal forces from the sun cause each protoplanet to become elongated with the long axis toward the sun, and consequently to rotate in the same length of time and in the same direction that it revolves about the sun (Figure 8-3). But as the protoplanet contracts under its own gravitational influence it rotates more rapidly, since its angular momentum must remain constant so long as no outside forces are acting upon it. Thus we see that the planets should rotate in the same direction in which they revolve. Similarly, the satellites form from the nebular disk about each protoplanet and thus revolve and rotate in the same direction that the planets rotate.

At first, this explanation of satellite formation does not appear to account for the rings of Saturn. However, E. Roche has shown that the tidal forces existing within a distance of 2.4 times the radius of a planet (the so-called Roche limit) are so great as to prevent the formation of a satellite. The concept of the Roche limit may then help explain the existence of Saturn's rings: the material originally composing the nebular disk surrounding Saturn which was within that limit has remained without coalescing into a satellite.

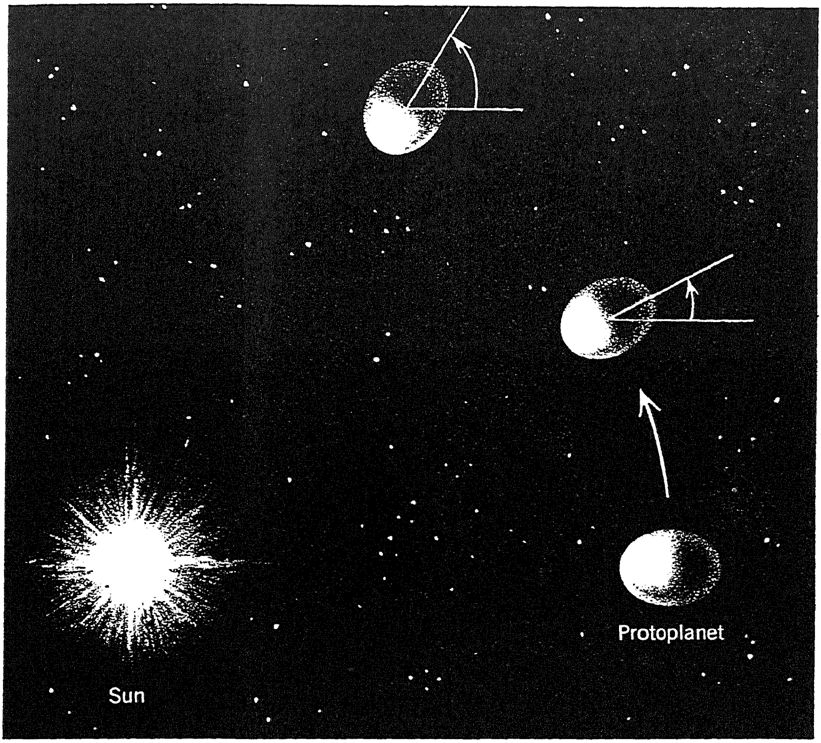


Figure 8-3 **The tidal bulge in a protoplanet must by its very nature line up with the sun and thus the protoplanet is given an initial rotation in the same direction that it revolves.**

The material outside the Roche limit formed the nine satellites of Saturn. We may recall that two centuries earlier Laplace also thought of Saturn's rings as evidence of a preexisting nebular disk.

Kuiper suggests that the entire process of formation can be discussed in four stages.

The first stage is the collapse of the original nebula into a proto-sun with a nebular disk rotating about it. The proto-sun in this stage would still be dark.

The second stage includes the formation of the protoplanets in the nebular disk. The disk would break up into two parts: the inner part from which the four terrestrial protoplanets form, and the outer part from which the four Jovian protoplanets form. In the region between these two parts the density of the nebular disk would not exceed the Roche density which is necessary for contraction to take place, and consequently the tidal forces imposed by the sun would prevent the formation of any one large body. It is conceivable that two bodies smaller than the planets would form, which upon colliding would break up and become the minor planets. The nebular disk beyond the Jovian protoplanets would also have a low density which would result in the formation of comets.

The third stage of formation brings the sun from darkness to a brightness slightly less than it has today. This intense radiation would ionize gases in the nebular disk relatively close to the sun. The interaction between this rotating, ionized disk and the sun's magnetic field would quite adequately account for the sun's loss of angular momentum and its consequential longer period of rotation.

The fourth and final stage in the formation process is a further reaction between the nebular disk and the radiation from the newly brightened sun. Just as comets' tails are forced away from the sun by radiation and particle pressure, so are most of the gases in the nebular disk. Only those gases that are tightly held by the gravitational field of a protoplanet would remain in the vicinity of the sun; the rest of the material (actually the vast majority of it) would be ejected from the solar system back into interstellar space. Those planets closest to the sun and thus subject to more intense solar radiation and particle pressure would lose more material than would the four Jovian planets. Since the ejected material would be composed largely of lighter substances (mostly hydrogen) this would account for the greater density of the terrestrial planets.

According to this hypothesis, the comets formed beyond Neptune, and each grew into a ball of frozen gases about a mile in diameter with silicate and metallic particles embedded in them. This, it may be recalled, is precisely what we think comets to be. It is quite likely that there are perhaps billions of comets forming a sort of halo about the sun at distances beyond the orbit of Neptune. Their unexpected trips into the inner parts of the solar system may be accounted for by the fact that Pluto, with its eccentric orbit, may travel through this region of comets, diverting some of them into orbits that bring them close to the sun. It is not impossible that other planets and even the nearest stars may also perturb some of the comets into eccentric orbits.

Nuclear Physics Helps

Recent studies of radioactivity and studies in the realm of nuclear physics have indicated new avenues of approach in our search for the manner in which our solar system came into being. There is every reason to believe that during the early stages of the solar system an isotope of iodine, iodine 129, existed as part of the primordial material from which the planets were created. The amount of iodine 129 created can be estimated from principles of nuclear physics which predict the ratio of iodine 129 created to that of iodine 127, common everyday stable iodine. But iodine 129 decays radioactively into xenon 129 with the relatively short half-life of 17 million years. If there was any iodine 129 left when any solid objects—presumably meteorites—formed they would have trapped the xenon 129 formed by this radioactive decay process. Therefore by measuring the amount of xenon 129 now in the meteorites one can estimate the amount of iodine 129 present at the time of their formation. Furthermore, by comparing the amount of iodine 129 present when the meteorites formed with the amount present in the primordial matter, John H. Reynolds of the University of California at Berkeley, was able (by knowing the half-life of iodine 129) to calculate the time interval between the formation of iodine 129 in the primordial matter and the condensation of the meteorites. His results indicate that only 100 million years had transpired! If the meteorites formed 4.6 billion years ago, iodine 129 formed 4.7 billion years ago.

It is interesting to note that in these studies by Reynolds and his student Craig M. Merrihue that there is *more* iodine 129 in the chondrules (see p. 161ff) than in the other parts of the meteorites. Consequently, it can only be concluded that the amount of iodine 129 had decreased in the nebular disk and the resulting xenon 129 dissipated out into space after the chondrules had formed and before the meteorites formed with the chondrules as inclusions. The chondrules must have solidified first. They are, it appears, the oldest solid objects in this solar system—they are real fossils of the nebular disk about the forming sun.

The fact that only 100 million years transpired between the formation of iodine 129 and the solidification of the chondrules changes our outlook on the formation of the elements. It has always been assumed—and with good reason—that the cloud from which the sun and its nebular disk formed was composed principally of hydrogen with very few heavy elements. Now it seems clear that those heavier elements (including iodine 129) were created on the surface of the newly forming sun by intense bombardment of high-energy protons.

Further research along these lines suggests that after the chondrules had formed, bodies coalesced which had an average diameter of about

12 meters (40 ft). These bodies formed after most of the hydrogen had been expelled from the inner regions of the solar system. It seems that the remaining hydrogen was trapped by the formation of water and ice. It is presumed that water was a part of these 40-ft bodies. If the planets formed from these objects, water must have become trapped inside the Earth.

Water in the material from which the Earth was formed would explain the geological theory that the waters of the oceans derived from water vapor expelled by volcanoes. Most of the "smoke" issuing from an erupting volcano is water; and it is this water which forms our oceans. Certainly there has been a lot of volcanic activity on the Earth since it formed.

Just how the planets formed from these bodies with a diameter of 40 ft—if indeed they did—remains to be seen, and this is not the only unanswered question in the study of early history of the solar system. But these new studies do indicate a change in the astronomers' approach to the problem. It is possible now—with the help of observations and theories based on work of the nuclear physicist and the geochemist—to approach the problem quantitatively. Measurements of time can be made. Studies of crystallography have made clear the conditions present when crystals were formed and any major changes in those crystals. As soon as the quantitative approach can be used, real progress toward understanding can be made.

Basic Vocabulary for Subsequent Reading

Accretion
Protoplanet
Radioactive decay
Radioactive half-life

For Further Reading

Hurley, Patrick M., *How Old is the Earth?* Doubleday Anchor, New York, 1959.

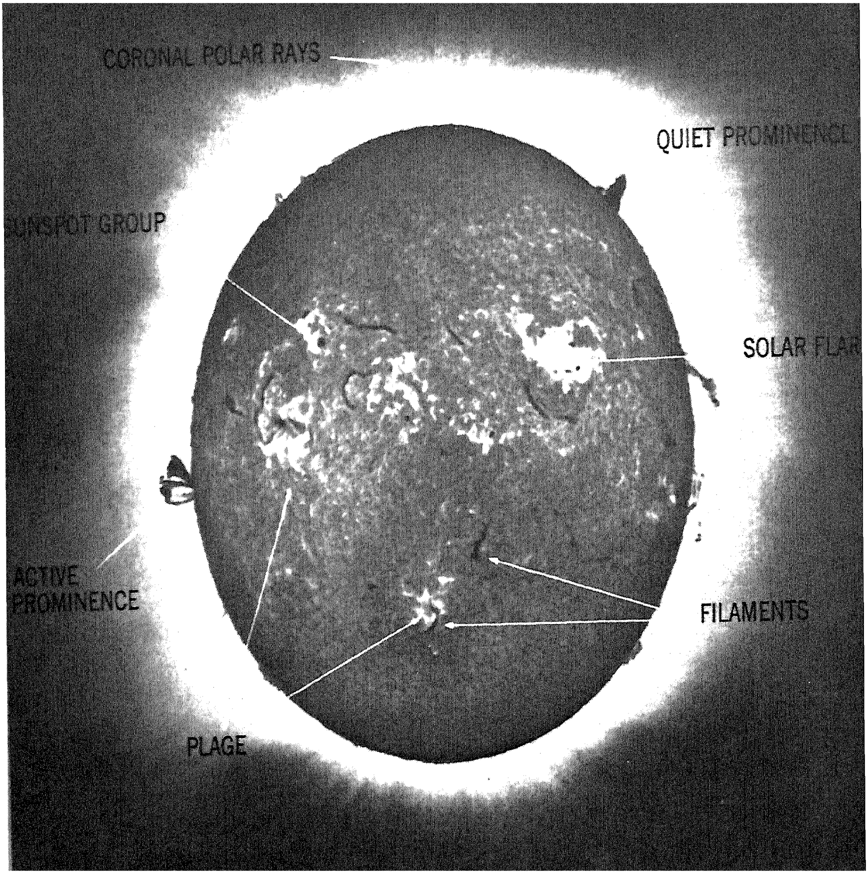
Gamow, George, *Biography of the Earth,* The Viking Press, New York, 1959.

Struve, Otto, and V. Zebergs, *Astronomy in the 20th Century,* Chapter IX, Crowell, Collier and Macmillan, New York, 1962.

Brown, Harrison, "The Age of the Solar System," *Scientific American,* p. 80, April 1957.

Huan, Su-Shu, "Recent Findings About Early Solar System History," *Sky and Telescope,* p. 13, July 1964.

"Oldest Solids," *Scientific American,* p. 72, March 1963.



Composite of two photographs of the sun, one of the disk in hydrogen light and the other of the limb with the disk masked out by an artificial eclipse. (Lockheed Aircraft Corporation)

Chapter 9 *The Sun*

At the center of the solar system resides its largest and most important member, an average star which we call the sun. The fact that it is relatively close makes possible fairly detailed observations that can tell us something about the other more distant stars. But of more immediate importance to us is the sun's influence on the Earth. Not only is it our source of heat and light, but it has other effects on the Earth that, although less noticeable, are interesting none the less.

The sun is a huge ball of gas* 864,000 miles in diameter, with a mass 333,000 times that of the Earth. But since its volume is 1,300,000 times that of the Earth, its density, 1.41 times that of water, is less than the Earth's density of 5.5, about the same as that of the Jovian planets. The force of gravity operating on its surface is 28 times that on the surface of the Earth. Therefore, if a man weighing 180 lb on the Earth could stand on its surface, he would weigh 5,040 lb. Such a large force of gravity also means that the escape velocity on the sun is very high, about 386 miles per second as compared with 7 miles per second for the Earth.

---

* This gas is quite unlike any found on Earth because its temperature, pressure, and density are so much higher than gases in our atmosphere. Yet the gases in the sun obey the gas laws met with on Earth.

The portion of the sun that we see in the sky, its visible surface, is called the *photosphere* (sphere of light). This surface is gaseous rather than solid like the surface of the Earth. The photosphere is visible because it is the sun's outermost region, where the pressure is great enough and the temperature high enough to emit a continuous spectrum. The photosphere is therefore opaque and hides the solar interior from our view. There is no abrupt change as there is between the Earth's surface and its atmosphere, but there is a more gradual change in density and temperature.

Above the photosphere there are two clearly differentiated regions of gas. Immediately above the photosphere and extending for an estimated 5,000 to 10,000 miles is the chromosphere. The chromosphere can be seen during a total solar eclipse just seconds before and after totality. During these fleeting moments, the moon covers the more intense photosphere and permits observation of the reddish chromosphere. In fact its reddish color is the origin of its name which means the "colored sphere." The gases of the chromosphere cause the absorption lines which show up in the solar spectrum by selectively absorbing light emitted by the photosphere. Accordingly, the chromosphere must be at a lower pressure and temperature than the photosphere which emits the background continuous spectrum.

From the chromosphere the corona extends for millions and millions of miles as a very tenuous gas. There is good evidence to conclude that many of the planets, including the Earth, actually revolve within the limits of the corona. The corona is best seen during a total solar eclipse, although both the chromosphere and the inner corona can be seen by optical means which simulate a solar eclipse.

Radiation Laws

One of the first questions that must be asked about the sun concerns the light and heat that it radiates into space. This amount of energy has been determined by measuring its heating effect at the Earth's surface and making allowance for the energy absorbed by the Earth's atmosphere. When this absorption is taken into account, we find that the Earth receives 2.0 calories per square cm each minute.* But at the distance of the Earth from the sun this 1 square cm represents a very small fraction of the total energy output of the sun. To find the total output, we must determine how many square centimeters there are on the surface of a sphere with a radius equal to the radius of the Earth's orbit. We can then multiply the amount of energy passing through 1 square cm by the total number of square centimeters on that sphere. The surface area of this large sphere is 2.9×10^{27} square cm. Consequently, the total output of radiation passing through this sphere must be $2.0 \times 2.9 \times 10^{27} = 5.8 \times 10^{27}$ calories per minute! This must be the total output of energy from the sun, which when translated into power rating is 5×10^{33} horsepower!

Not only is the total amount of energy radiated by the sun important to the astronomer, but the amount of energy radiated at each wavelength is of vital interest, for it enables him to calculate the temperature of the photosphere by three different methods using, in turn, the Stefan-Boltzmann law, Wien's law, and Planck's law. Each of these three laws describes certain aspects of the *spectral energy curve*, that is, the graph showing how the intensity (amount of energy per unit wavelength) of radiation varies with wavelength (Figure 9-1).

The area under this curve is given by the Stefan-Boltzmann law and represents the amount of energy radiated per unit surface area. According to this law, the total amount of energy radiated is not only proportional to the area A of the surface; it is also proportional to the fourth power of the temperature T:

$$E \propto AT^4$$

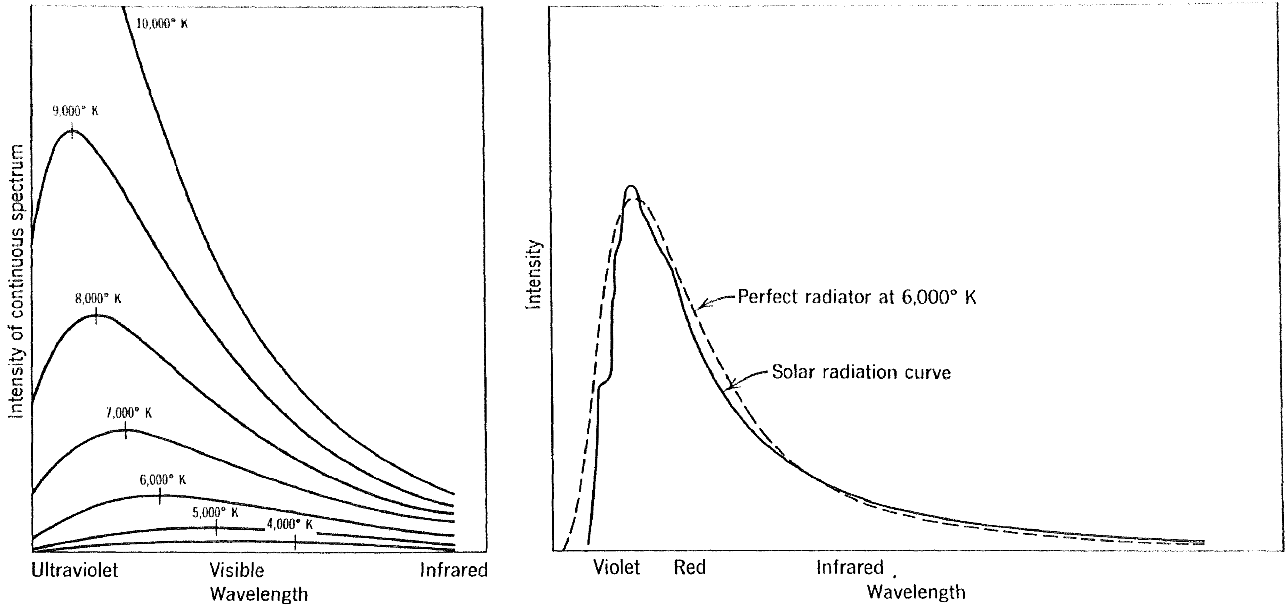
or written as an equation

$$E = \sigma AT^4$$

where E is the total amount of energy radiated each second. The proportionality is changed into an equation by the insertion of a constant of proportionality σ (Greek small sigma). The value of σ is

---

* The calorie, a unit of heat, is the amount of heat required to raise the temperature of 1 cubic cm of water 1°C. The calorie referred to in dietetics is actually the kilocalorie, or 1,000 calories.



(*a*) (*b*)

Figure 9-1 (*a*) Radiation curves for various temperatures. As the temperature increases the amount of radiation increases (area under each curve) and the wavelength of maximum radiation (the peak of the curve) shifts to the shorter wavelength. (*b*) The radiation curve of the sun matches more closely the radiation curve of a black body at 6,000° K than that at any other temperature.

5.67×10^{-5} if T is in degrees Kelvin* and the area is expressed in square centimeters. For a given temperature, the greater the surface area, the more energy radiated. For a given surface area, the higher the temperature the more energy radiated. But we know the surface area of the sun, and we have just calculated the amount of energy it radiates; from these we can calculate the sun's surface temperature as 5,770°K.

In order to check this value, we may use Wien's law: as the temperature of a radiating body increases the peak of the spectral energy curve (the wavelength of maximum radiation, λ_{peak}) shifts to the shorter wavelength $\left(\lambda_{\text{peak}} = \dfrac{2.9 \times 10^{7}\ \text{Å}}{T}\right)$. It may be recalled from Chapter 3 that the shorter wavelengths of light have more energy than the longer wavelengths. Since a higher temperature represents more energy we would expect a hot object to radiate energy of shorter wavelengths than a cooler one. We must, therefore, determine that particular wavelength of light for which the energy radiated is stronger that at any other wavelength.

Wien's law is used to determine the temperature of a furnace by noting the color of its inside. As the furnace heats up it begins to glow with a dull red color, but as the temperature increases its glow becomes first more orange, then more yellow. That is, as the temperature increases, more light of shorter wavelengths (yellow, green, blue) is emitted and its color begins to approach "white heat."

This change of color with increasing temperature is also seen in the stars; for a cool star (cool, in terms of stellar temperatures, means about 3,000°K) appears red, whereas a hot star (about 20,000°K) will be blue. For if an object becomes sufficiently hot, it will go beyond "white heat" to appear ultimately blue.

The sun does not even reach a "white heat" but appears slightly yellow despite the fact that the peak of its radiation curve is in the blue region of the spectrum at 4,750 Å (Figure 9-1*b*). According to Wien's law this corresponds to a temperature of 6,100°K, slightly more than that determined by the Stefan-Boltzmann law.

---

* Scientists and the non-English speaking world use a temperature scale called Celsius (centigrade) designed so that water freezes at 0°C and boils at 100°C. There is another temperature scale also used by scientists called absolute or Kelvin. Zero on the Kelvin scale is absolute zero, that is, that temperature at which atomic motion supposedly stops. According to present theory and practice nothing can get as cold as 0°K (−273°C or −460°F). Water freezes at 273°K and boils at 373°K.

Since the wavelength of maximum radiation, 4,750 Å, is in the blue region of the spectrum, why doesn't the sun appear blue? There are two reasons for this. The first is that the Earth's atmosphere scatters blue light better than any other color in the visible part of the spectrum. The light that comes through unscattered must appear less blue. The second reason is that even though the peak of the energy curve is in the blue, the sun still radiates much green, yellow, orange, and red light, which combines with the blue to produce a near-white.

Planck's law actually embodies the two radiation laws just discussed, for it gives us the shape of the spectral energy curve and by giving us the shape it automatically gives the area underneath the curve as well as the wavelength of the peak of the curve. A perfect radiator at a temperature of 6,000°K would radiate light according to the dotted curve (Figure 9-1*b*). The sun's radiation curve indicates a temperature close to 6,000°K.

The reason for the difference between the temperature values derived by means of the three laws of radiation lies in the fact that the sun is not a perfect radiator. A perfect radiator is one which, if it were cold, would absorb all the light incident upon it and would therefore appear completely black. Such a perfect radiator is called a *black body,* and when incandescent it would radiate light more efficiently than any material substance and in strict accordance with Planck's law.

The Solar Spectrum

The studies of the solar spectrum now exceed the bounds of the optical spectrograph. The sun is being studied in nearly all wavelengths of the electromagnetic spectrum. Both spectrographic studies and direct photographs of the sun are being made in the X-ray and ultraviolet region by rockets and satellites which operate above our shielding atmosphere. Optical studies (including the near-ultraviolet and infrared) can be made with ground-based telescopes. Radio astronomers are carrying on an intense study of the solar radiation in the long wavelengths of the electromagnetic spectrum.

Spectral studies in the optical region reveal some 26,000 spectral lines! Each of these lines must result from a particular element or chemical compound, and only 18,000 lines have been identified. Of the more than 100 elements on the Earth, only 67 have been observed in the sun. The others may be exceedingly rare or have spectral lines in regions of the electromagnetic spectrum not well studied yet.

It is not easy to identify spectral lines; temperature and pressure conditions on the sun are too different from those prevailing on the Earth. We can identify many lines such as those of hydrogen, iron, ionized calcium, helium, and others. But even for some of these elements, such as iron, a spectrum obtained in a terrestrial laboratory will not contain all the iron lines that are found in the solar spectrum.

The identification of spectral lines is facilitated by a consideration of their *intensity*. The intensity of an emission line is, of course, the amount of light in that line, whereas the intensity of an absorption line is the amount of light subtracted from the continuous background spectrum. The intensity of a spectral line resulting from a given type of atom in a given state of excitation is a direct indication of the number of atoms in that state. Thus, if many atoms of the same kind experience the same transition the resulting line will be more intense than if relatively few atoms experience that transition. For example, as more atoms of iron become ionized the resulting lines become more intense, whereas the lines resulting from neutral atoms of iron become less intense.

These principles help us in the theoretical studies of energy levels which must be made before we can identify the lines not seen in terrestrial sources. By means of such studies we can determine which transitions are most likely to occur under conditions on the sun; the greater the probability, the more intense the line should be.

The two main factors influencing electron transitions are the temperature and the pressure of a gas which is either emitting a bright-line spectrum or causing absorption lines in a continuous spectrum. To illustrate this, let us first consider a gas at moderately low pressure whose temperature increases from a comparative coolness (let us say 3,000°K) to a high temperature. When the gas is cool only those transitions that require a little energy will occur and therefore we would expect most of the lines to appear in the red region of the spectrum. In a gas at this temperature it is even possible for some atoms to remain bound to each other, forming molecules. Consequently we would see some molecular bands, such as those of titanium oxide.

As the gas is heated the molecules begin to break up, and transitions occur that require more energy. Consequently the spectral lines (either bright or absorption, as the conditions of the gas dictate) toward the violet end of the spectrum begin to appear and become stronger. Furthermore, some of the atoms begin to lose their outer electrons and as ions produce spectral lines, such as ionized calcium. Lastly, when the gas is hot more atoms have become ionized, and the many transitions which require that energy produce spectral lines deep into the ultraviolet.

If we keep this gas at high temperature and only moderately low pressure, the atoms and free electrons are relatively close together; consequently the ions will be able to capture electrons fairly easily and may not remain ions very long. If the temperature is maintained but the pressure reduced still more, the ions and free electrons will be farther apart and a given atom, if ionized, will remain ionized for a longer period of time. Therefore, of two gases at the same temperature, the one with the lower pressure will show more intense lines of ionized atoms.

A striking example of the identification of lines unseen in terrestrial sources presents itself in the so-called "coronal" lines. These lines in the spectrum of the solar corona remained unidentified for years before a theoretical study of the sort discussed previously showed that they result from transitions in iron atoms that have been ionized no less than thirteen times. Since an atom of iron, when electrically neutral, contains 26 electrons, some of the atoms in the solar corona have lost half of their electrons. Such extensive ionization indicates not only a very high temperature but also a very low pressure. Similar studies identified helium in the sun before it was known to exist on the Earth. The name (from *helios*) gives due credit to the locale of its discovery.

We have seen that the solar absorption spectrum (not including the coronal spectrum which is seen at the time of total solar eclipse) results from the gases in the chromosphere. Studies of this spectrum indicate that the bottom of the chromosphere has a temperature of about 4,500°K and a pressure only about 0.1 that of the Earth's atmosphere at its surface. But gases at this temperature and pressure ought to emit a bright-line spectrum rather than cause an absorption-line spectrum. The chromosphere causes an absorption-line spectrum because it is normally seen in front of the photosphere, which is hotter (6,000°K) and therefore emits more radiant energy at each wavelength.

During a total solar eclipse, however, when the moon covers the rest of the sun, the chromosphere can be seen projected against

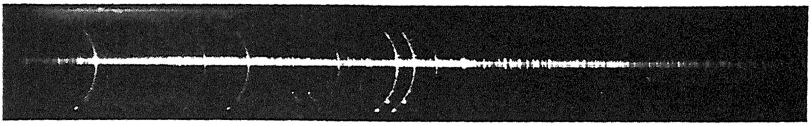


Figure 9-2 Flash spectrum of total solar eclipse of August 31, 1932, in Fryeburg, Maine. (Lick Observatory)

a dark sky; it then emits a bright-line spectrum. This bright-line spectrum occurs within less than a minute before and after totality, when the visible portion of the chromosphere is shaped like a very thin crescent. If the light from this crescent is allowed to pass through a spectrograph that has no slit, it produces an effect similar to that produced by a slit, except that the lines appear not as lines but as thin crescents (Figure 9-2). Such a spectrogram is called a *flash spectrum* because the appearance of the chromosphere just before and after totality is a sudden flash.

The flash spectrum is very interesting because it gives us the opportunity to study the distribution of elements in the solar atmosphere. There are a few lines that form longer crescents than others. These are the lines of calcium that have been ionized once and of hydrogen. Since their crescents are longer, these gases must reach higher into the atmosphere than the other gases. The visible hydrogen spectrum is composed of one very bright line in the red ($H\alpha$) and other lines of the Balmer series in the blue and violet. The two bright lines of ionized calcium (called H and K, following Fraunhofer's designation) are in the violet. The combination of these lines causes the reddish color of the chromosphere during a total solar eclipse.

The Sun's Interior

The sun's interior is of utmost importance in understanding the more obvious solar features; it is the furnace that keeps the sun "burning." The interior, however, must be more than a furnace, for the sun is too hot to burn—fire as we know it cannot exist there because the sun is too hot for the common chemical reactions to take place. The source of light and heat in the sun is the energy derived from nuclear reactions. These nuclear reactions take place deep inside the core of the sun where the temperature is hot enough (about 15,000,000°K) to force atoms together so fast that instead of rebounding some of them combine to form new atomic nuclei.

There are two principal nuclear reactions that could occur in the sun's core—the *proton-proton reaction* and the *carbon cycle.* The proton-proton reaction is not only simpler but also is the most likely explanation for the sun's source of energy, since the temperature of the sun's core is only moderately hot as stellar interiors go. At 15,000,000°K, two protons (nuclei of hydrogen atoms) may combine, if they collide, to form a deuterium nucleus in accordance with the following equation:

$${}_1H^1 + {}_1H^1 \rightarrow {}_1H^2 + {}_1e^0$$

The H is the chemical symbol for hydrogen. The number in front of and below each letter gives the *atomic charge* (the number of protons in the nucleus). The number following and placed above the chemical symbol represents the *atomic mass number* (the number of protons and neutrons in the nucleus). The ${}_1e^0$ is a *positron* that takes away the remaining positive charge and some energy. It does not have a long life, however, for it quickly combines with a nearby electron to form two gamma-ray photons. The ${}_1H^2$ is an isotope of hydrogen called *deuterium.* It behaves chemically like ordinary hydrogen, but has twice the mass. The deuterium nucleus will, upon collision, combine with another proton to form an isotope of helium whose mass is three:

$${}_1H^2 + {}_1H^1 \rightarrow {}_2He^3 + \text{gamma ray}$$

One helium-3 nucleus may react with another to form helium-4 along with two protons:

$${}_2He^3 + {}_2He^3 \rightarrow {}_2He^4 + 2\,{}_1H^1$$

A total of six protons go into this reaction, and two protons come out plus one helium-4 nucleus, with the net result that four hydrogen atoms are converted into one helium atom.

The carbon cycle, on the other hand, relies on nuclear reactions involving heavier elements, so a higher temperature is required before it can take place. It therefore plays a more important role in hotter stars, although it may contribute toward energy generation in the sun as well. The following six equations represent the various nuclear reactions involved in the carbon cycle:

(1) $${}_6C^{12} + {}_1H^1 \rightarrow {}_7N^{13} + \text{gamma ray}$$

(2) $${}_7N^{13} \rightarrow {}_6C^{13} + {}_1e^0$$

(3) $${}_6C^{13} + {}_1H^1 \rightarrow {}_7N^{14} + \text{gamma ray}$$

(4) $_{7}N^{14} + {}_{1}H^{1} \rightarrow {}_{8}O^{15}$ + gamma ray

(5) $_{8}O^{15} \rightarrow {}_{7}N^{15} + {}_{1}e^{0}$

(6) $_{7}N^{15} + {}_{1}H^{1} \rightarrow {}_{6}C^{12} + {}_{2}He^{4}$

Carbon 12 combines with a proton to yield nitrogen 13 plus a gamma ray. Nitrogen 13 is radioactive and decays into carbon 13 plus one positron. Carbon 13 combines with a second proton to form nitrogen 14 and a gamma ray. Nitrogen 14 combines with a third proton to yield oxygen 15 (plus a gamma ray), which is radioactive and decays into nitrogen 15 and a positron. Nitrogen 15 combines with a fourth proton to give carbon 12 once more, plus a helium atom.

As in the proton-proton reaction, the net result is that four hydrogen atoms combine to form one helium atom. The carbon used in the first equation is replenished in the last; so it is not used up by the process but acts as a catalyst.

The energy from both these sets of reactions results from the fact that four protons have a total but uncombined mass of 4.0291, whereas one helium nucleus has a mass of 4.0015. Thus, 0.0276 mass units are lost in each reaction. But according to the law of conservation of matter and energy, mass cannot be destroyed; any loss of mass must be converted into energy according to the equation formulated by Einstein, $E = mc^2$, where m represents the amount of mass lost in the reaction and c is the velocity of light. We previously calculated the amount of energy radiated by the sun to be 5.8×10^{27} calories each minute. This is equal to 4.0×10^{33} ergs per second (more convenient units) which can be substituted into the equation as E. The velocity of light is $c = 3 \times 10^{10}$ cm per second. By solving the equation for the mass m, we find that the sun loses 4.5×10^{12} grams of mass each *second*. On the Earth, that amount of matter would weigh nearly 5×10^{6} tons!

But can the sun continue losing mass at this rate for 4.6×10^{9} years and still have some left? At this rate the sun would lose 6.5×10^{29} grams in that time span. But the mass of the sun is 2.0×10^{33} grams, so the sun would lose only 0.03% of its present mass! By way of comparison, if a 125-lb lady on a diet were to lose 0.03% of her weight, she would lose the insignificant amount of ½ ounce!

The rate of both these reactions depends not only on the temperature but also on how closely the atoms are packed together. From theoretical studies, it seems very probable that the core of the sun

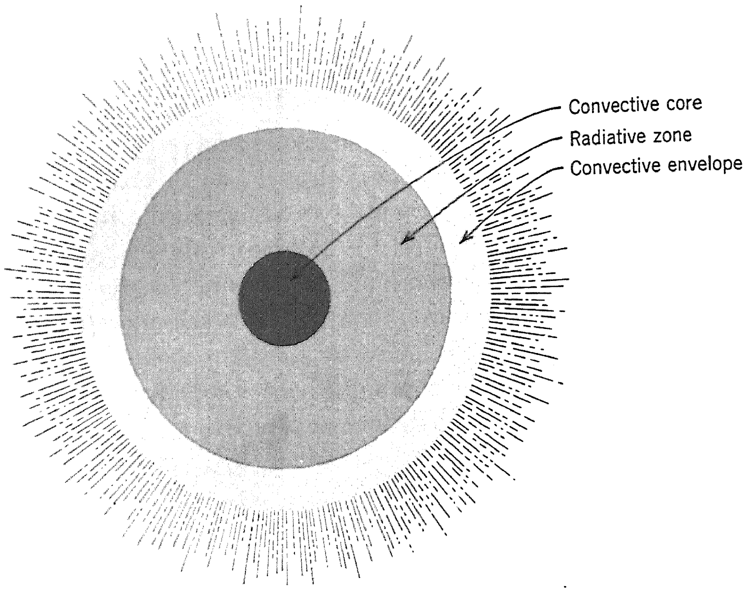


Figure 9-3 A cross section of the sun.

is under a pressure that is equal to about 3×10^{12} lb per square in. (2×10^{11} atmospheres). The resulting density, about 110 times that of water, means that 1 cubic ft of material from the sun's core would weigh about 6,850 lb if placed on the Earth!

The energy that is generated deep in the interior of the sun must make its way to the surface in order to radiate out into space; if the energy could not get out, the sun would get hotter and hotter and consequently would expand. Once the energy leaves the core of the sun, it travels outward through the *radiative zone* in the form of X-rays and gamma rays until it reaches a region called the *convective envelope* (Figure 9-3) which is quite opaque to X-rays. The energy must travel in the form of convective currents similar to the air above a hot road in summertime, but on a much grander scale. Because such a vast amount of energy strikes the convective envelope from below, these currents become turbulent; the result is the turmoil of the photosphere.

Photospheric Activity

The photosphere's mottled appearance indicates regions of different temperature called *granules* (Figure 9-4). The bright granules are

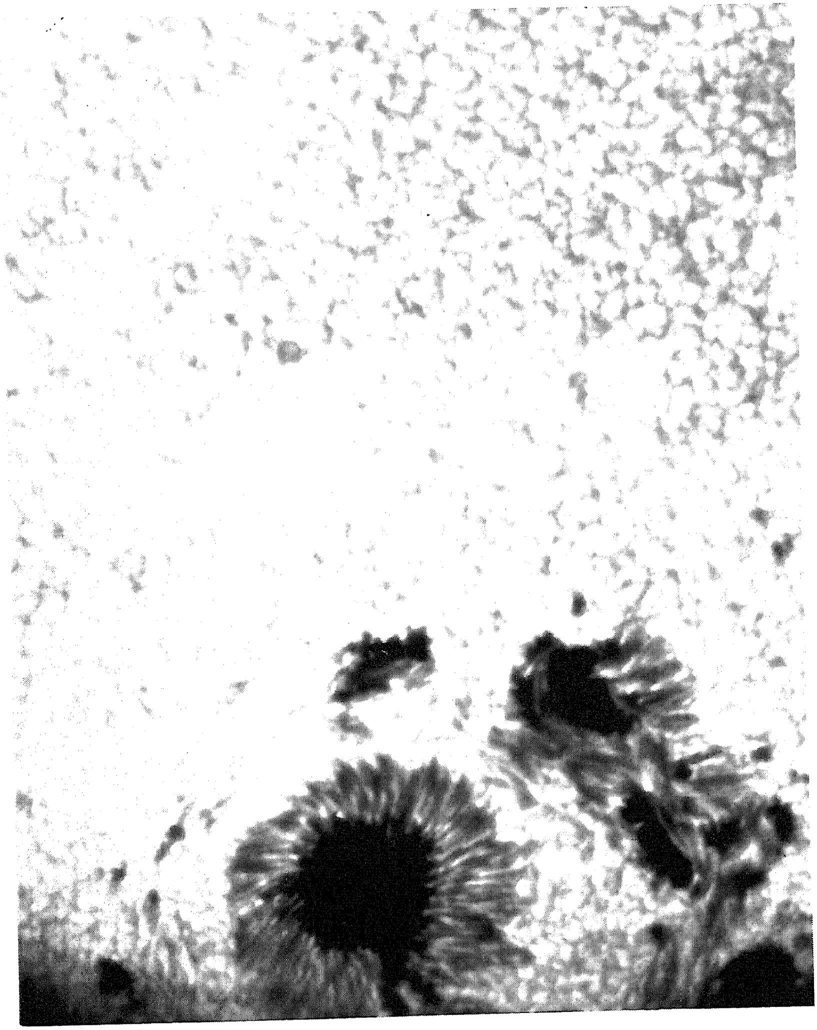


Figure 9-4 A sunspot group and solar granulation as photographed from a balloon at an altitude of 80,000 ft on August 17, 1959. This photograph was taken as part of Project Stratosphere of Princeton University, sponsored jointly by the Office of Navel Research and the National Science Foundation. (Courtesy of M. Schwarzschild)

several hundred miles in diameter, last for only a few minutes, and are hotter than the darker regions. They result from a parcel of gas which is heated from below, rises, and appears bright. In rising, the gas expands, radiates energy, cools, and sinks back into the lower levels to become part of the darker portion of the photosphere. There the gases are again heated and rise once more as a bright granule.

The turbulent nature of the photosphere is revealed not only indirectly by these rapid changes in the granulation but spectroscopically as well, for spectral lines formed in a turbulent gas will be broader than lines formed in a quiet gas. Since atoms in a turbulent gas move both up and down, their motions cause simultaneous Doppler shifts to the red and blue. The spectral lines are therefore broadened in both directions. This is called *turbulent broadening.* Because it is evident in addition to and can be differentiated from pressure broadening (see p. 56), astronomers have learned that the gases in the chromosphere must move with velocities of about 1 mile per second. Such turbulence results from energy transported up from the sun's interior.

With the very high dispersion of the big solar telescopes, the turbulent broadening of the solar spectral lines can be resolved into irregularities in the lines (Figure 9-5). Immediately adjacent portions of the three spectral lines are shifted in different directions. The direction in which each portion is shifted, depends upon where in the lower chromosphere that portion of the line is formed. If it is formed in a rising column of gas, that portion of the line is shifted to the violet; if it is formed in a sinking parcel of gas, it is shifted toward the red. The same portion of each of the three lines is shifted in the same direction, for they were each formed in the same parcel of gas.

In our everyday life we consider the sun's outpouring of radiant energy very nearly constant; nevertheless it does manifest slight cyclical variations over a period of about 11 years. The most obvious evidence for this is the spotty appearance of the photosphere. Dark regions, called *sunspots*, appear from time to time. Some of the sunspots are small and normally do not last very long. Others are very large by terrestrial standards and may last as long as several months. Some spots are single and travel with the rotating sun as a lone spot. It is more common, however, for the spots to appear in groups (Figure 9-6). Each group is usually divided into two main parts—the

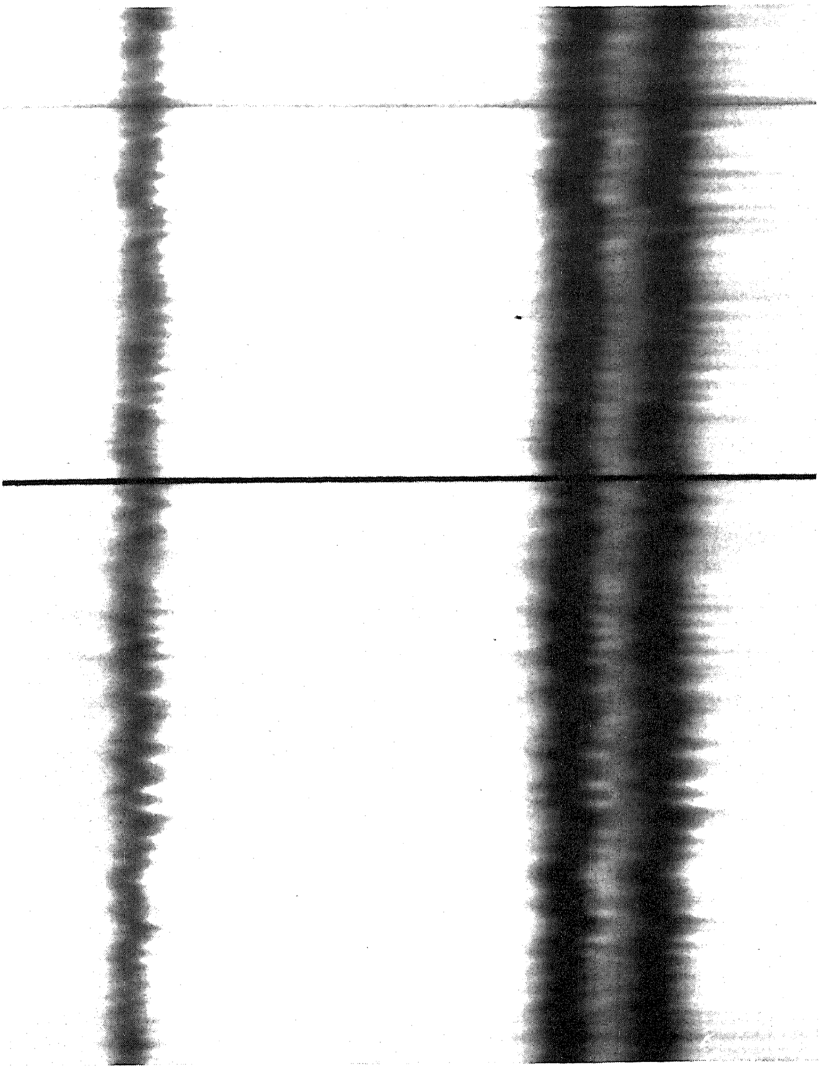


Figure 9-5 High resolution photographs of three solar spectral lines. The horizontal solid black line represents a black spot on the sun's surface, but is actually in the spectrograph. Actual locations on the sun's surface can be recognized by the horizontal white and grey streaks. The white streaks result from hot gases rising to the surface. The three lines and wavelengths are (left to right): iron, 5,187.9 Å; ionized titanium 5,188.7 Å; calcium 5,188.9 Å. (Sacramento Peak Observatories, Air Force Cambridge Research Laboratory)

preceding spot (preceding in the sense of solar rotation) and the one following.

Generally a group in the process of formation initially consists of two spots, which within a few days will tend to drift apart in longitude. The preceding spot is usually the larger, more regular one, and is less likely to undergo rapid changes. The following spot is the first to disappear; it breaks up into smaller spots that gradually decrease in size until, after an interval of days or weeks, they disappear. The preceding spot does not usually break up but merely decreases in size until after a period of weeks or even months it, too, disappears.

The center of a sunspot, being dark, is called the *umbra.* It is surrounded by the *penumbra,* which is brighter than the umbra but still not as bright as the photosphere. The spot is dark when compared with the photosphere which is indicative of a lower temperature; this fact is verified by spectrograms made of sunspots. In these spectrograms we find lines that are typical of a gas at 4,000°K. But a sunspot is not black in any sense of the word; if the rest of the sun were to be masked off we could see a large bright spot with the naked eye. It is dark only when compared with the photosphere, because its temperature is about 2,000°K less.

Since sunspots rotate with the sun, they give us a clue to the sun's period of rotation. But, as often happens, the clue leads to another problem: we find that this period is not the same for all latitudes. If a spot appears at the equator, its period of rotation will be about 25.1 days; if it appears at 15° north or south latitude, its period of rotation will be about 25.5 days; and if it appears at 30° latitude, its period of rotation will be about 26.5 days.

In order to verify this finding we can obtain the period of rotation by placing the limb of an image of the sun on the slit of the spectrograph and determine the Doppler shift. Spots are rare at distances of more than 35° north or south of the equator, and consequently the spectrographic method proves more useful because it can be used at much greater latitudes. The periods of rotation determined spectrographically, however, are not in complete accord with the periods determined from sunspots. The spectrograph tells us that the equator rotates in 24.6 days; 15° latitude in 25.4 days; 30° in 26.4 days; 60° in 31.0 days; and 75° latitude in 33.0 days. That a discrepancy should exist may be accounted for by the fact that whereas the spots yield a period of rotation of the photosphere, the spectrographic determina-

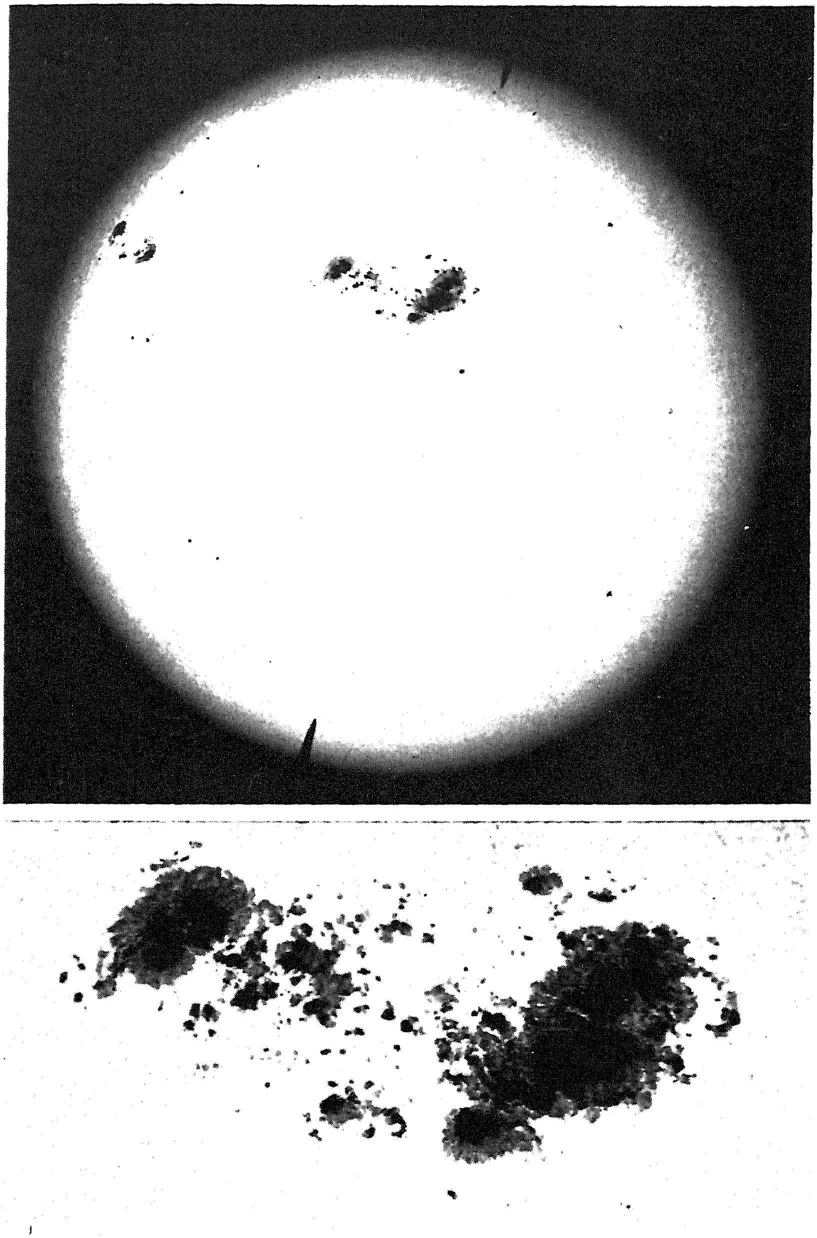


Figure 9-6 A large group of sunspots, April 7, 1949. (Mount Wilson and Palomar Observatories)

tion yields a period of rotation of the lower chromosphere where the absorption lines originate.

Sunspots were observed as far back as 28 B.C. by the Chinese who saw "flying birds" on the solar disk. Galileo observed them with his telescope in 1610, and they have been recorded with some reliability since about 1750. Not until 1851, however, was it recognized that the number of sunspots varies over the years. Figure 9-7 shows that the *sunspot cycle* is not regular; neither all the maximum years nor all the minimum years have the same number of spots. In fact, even the period of 11 years is not the same. In 1917, for example, the maximum year occurred with a sunspot number of 104; 11 years later, in 1928, there was another maximum year with a sunspot number of 78; only 9 years later, 1937, with a sunspot number of 114, was another maximum year.

The number of spots is not the only variation, however, for the latitude at which spots occur changes as the cycle progresses. A cycle begins during a minimum, at which time the spots in the new cycle appear at latitudes of about 30° north and south. A given spot maintains approximately the same latitude, but as the cycle progresses the succeeding spots appear ever closer to the equator, until by the time minimum again occurs the spots in that cycle appear at about 10° north and south of the equator. Often two cycles overlap; a few spots belonging to the old cycle may appear near the equator, while the spots of the next cycle are appearing at about 30° from the equator. Spots are rarely seen more than 35° from the equator, and only occasionally do they reach the equator. Figure 9-8 shows a distribution of sunspots according to latitude and years.

During the 11-year cycle, then, both the number of spots and their latitude vary. But there is yet another variation, magnetic in nature. Associated with each spot is a magnetic field that behaves as if a long bar magnet had been inserted into the spot with one end sticking out. The evidence for this is spectrographic: when a spectral line is formed in a strong magnetic field it will be split into several components (the Zeeman effect). The characteristics of this splitting depend on the strength of the magnetic field and the direction in which the light leaves it.

Studies of the magnetic fields associated with sunspots reveal that if a preceding spot in a group has the north pole of the magnetic field pointing out from the sun, the following spot will have the south pole sticking out. The spots in the other hemisphere (northern

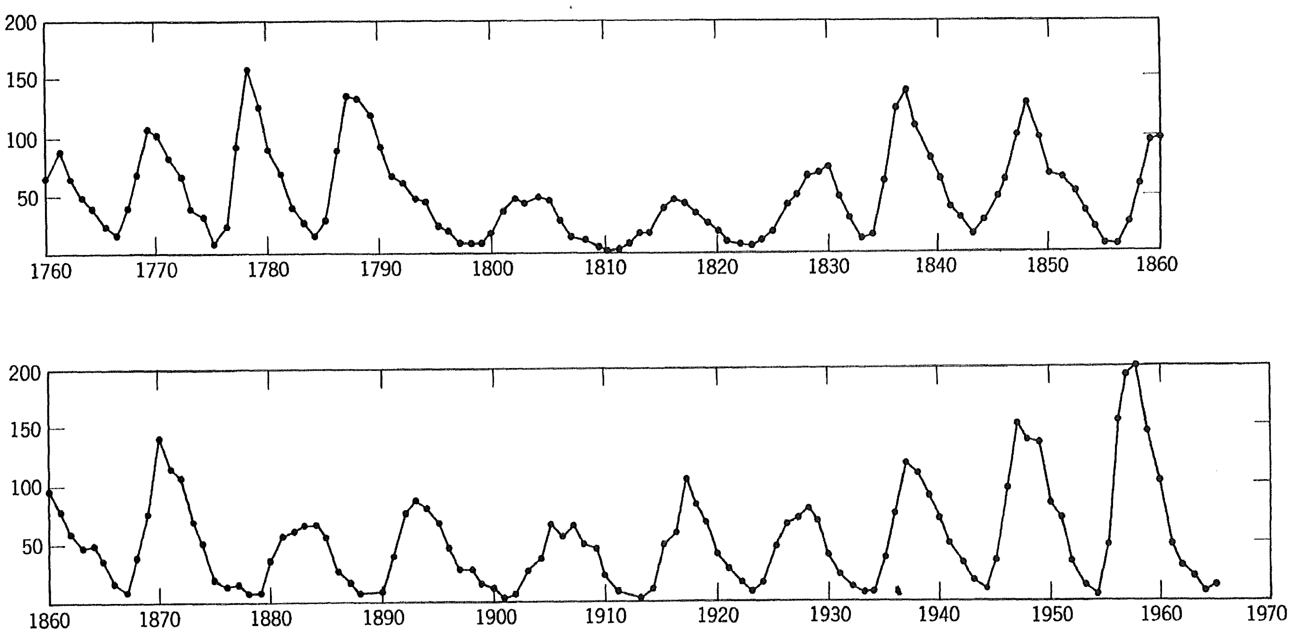


Figure 9-7 The sunspot cycle. The variations in the number of sunspots since 1760. The records before 1860 are not very reliable for they were not kept systematically.

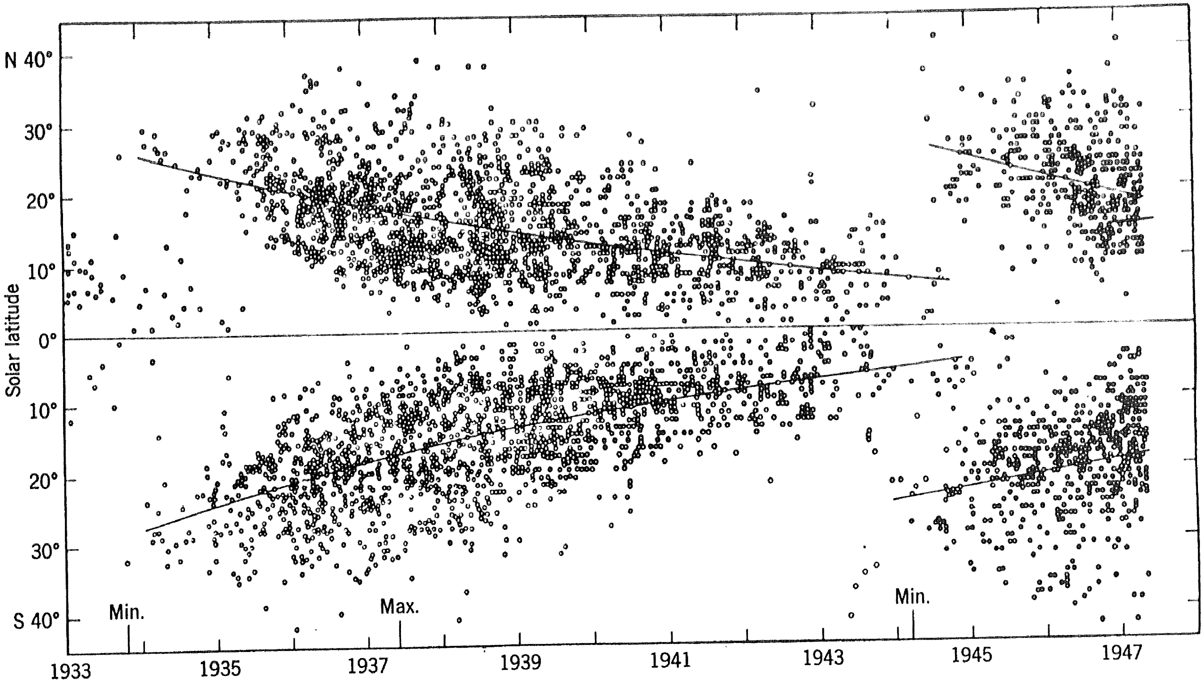


Figure 9-8 The distribution of sunspots during the 11-year cycle. At the beginning of each cycle the sunspots form at some distance from the equator. As the cycle progresses they form ever nearer the equator. (Courtesy of Giorgio Abetti, *The Sun*, The Macmillan Co., 1957)

or southern) will have the opposite magnetic polarity; the preceding spot will be a south magnetic pole and the following spot will be a north pole. During the next sunspot cycle, however, the polarity is reversed in both hemispheres. Any two succeeding cycles, then, are different in their magnetic aspects; consequently the cycle does not repeat itself in every respect until about 22 years have elapsed.

The reason for this magnetic variation is quite unknown; in fact, we do not know with certainty why sunspots have magnetic fields associated with them at all. But it is suspected that the magnetic field causes the spot, not the spot the magnetic field. The main problem is one of refrigeration! How can the magnetic field maintain the temperature 2,000° cooler than the surrounding photosphere?

We could learn more about the photosphere in general and sunspots in particular if we knew where the various spectral lines originated in the sunspots and in the chromosphere. Spectral lines are, after all, indicative of the conditions under which they form. Since some of the lines in the solar spectrum are rather broad and quite intense, it is possible to take a picture of the sun by the light of any one of these broad lines alone. (It should be recalled that even absorption lines are not black but contain some light that has been absorbed and re-emitted by the gas.) Special filters, both mechanical and optical, have been devised that eliminate all the solar spectrum except that spectral line of special interest. The resulting photograph, called a *spectroheliogram,* reveals the distribution of that particular gas whose spectral line has been allowed to pass through the filter. The three principal lines used in studying the sun's surface are $H\alpha$ of hydrogen and the H and K lines of ionized calcium, since these lines are sufficiently intense to yield a good picture, and broad enough to be isolated by filters.

Figure 9-9 shows four photographs of the sun, each taken at the same time: (*a*) was taken by the light of the entire visible spectrum, (*b*) was taken by the light of $H\alpha$ alone, and (*c*) was taken by the light of the K line of calcium. Photograph (*d*) was taken by the ultraviolet light of the Lyman Hα line. Since the Lyman α region of the spectrum is absorbed by our atmosphere, this latter photograph was obtained from above the atmosphere by a camera in an Aerobee-Hi rocket on March 13, 1959.

The brighter regions in each of the four photographs are regions of the sun that are at higher temperatures and consequently these regions emit more light per unit area than the rest of the surface.

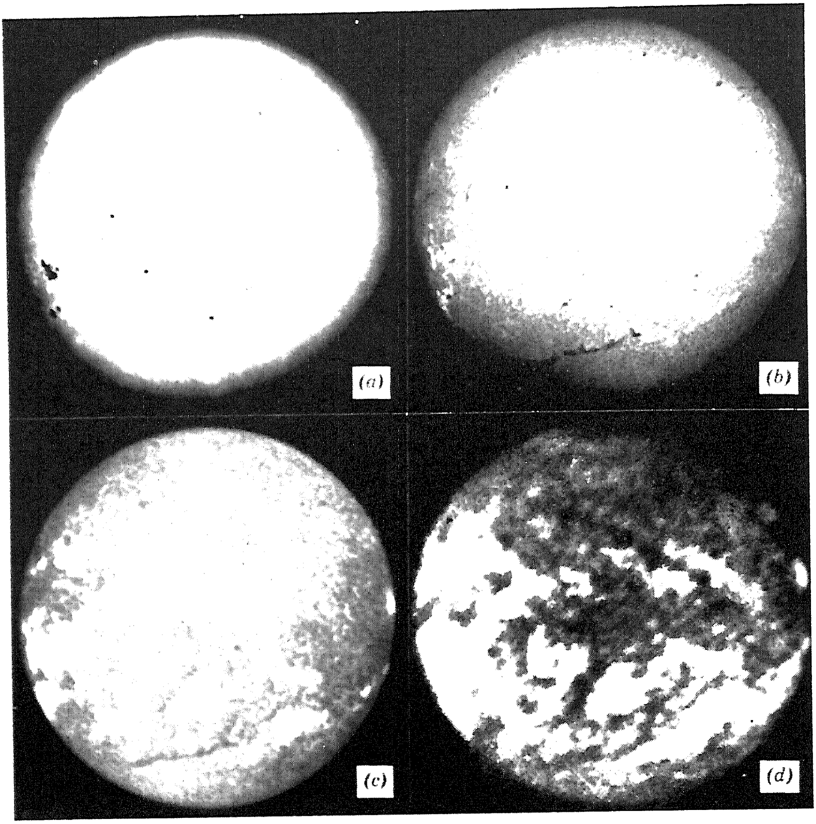


Figure 9-9 The sun photographed simultaneously by (*a*) white light; (*b*) the light of hydrogen; (*c*) the light of calcium; (*d*) the ultraviolet light of the hydrogen Lyman alpha line. (U.S. Navy)

Each brighter region is called a *plage* and is evidence of some disturbance on the sun. Sunspots are located in plages (Figure 9-9*a*).

When the sun is photographed by means of hydrogen light, the mottled appearance of photographs made by calcium light becomes less evident, and we see striations or streaks as in Figure 9-10*a*. Some of the streaks are bright, indicating a higher temperature, and nearly all of them appear to follow magnetic lines of force as iron filings do about a bar magnet. One other feature that is often visible in spectroheliograms of hydrogen light is the long *dark filaments*. These are gases above the surface that absorb light.

Another form of photospheric activity is the *solar flare*, a very hot and explosive region. A flare normally forms in a sunspot group, generally along the line dividing the opposite magnetic polarities (Figure

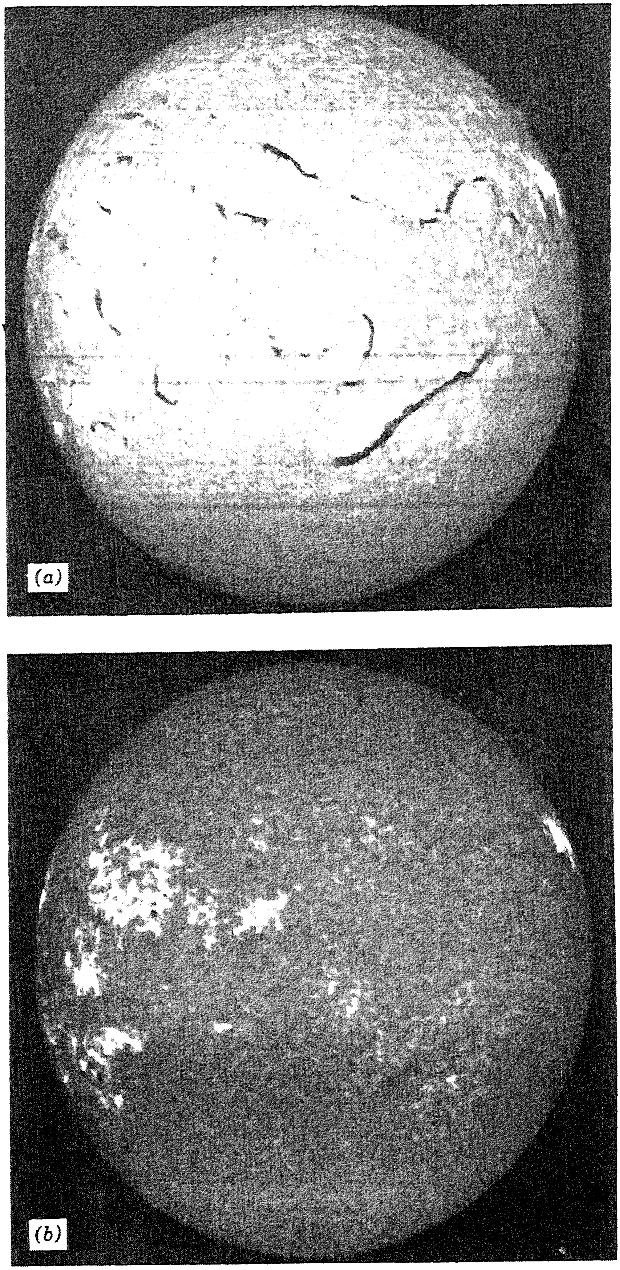
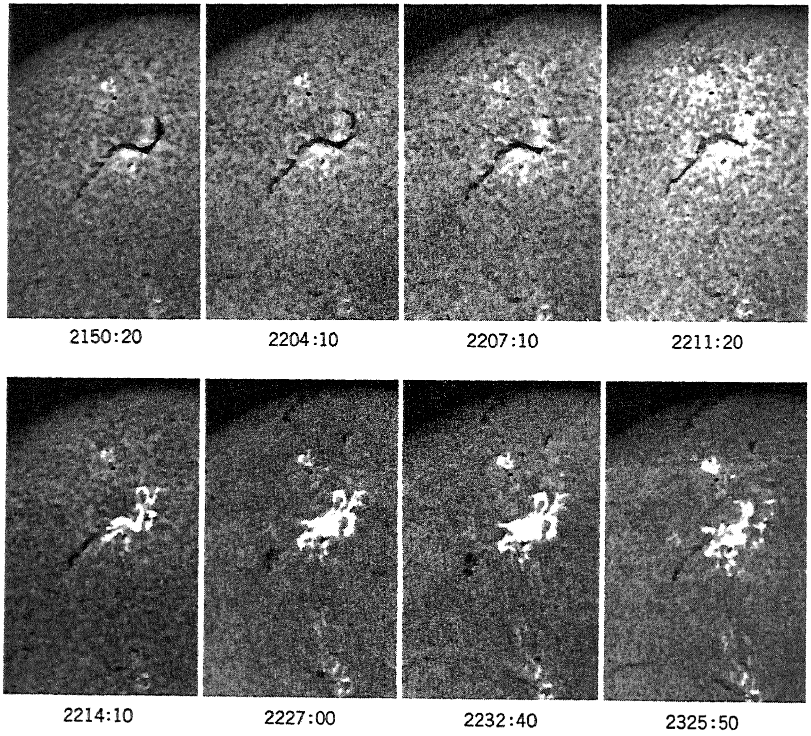


Figure 9-10 The sun photographed in (*a*) hydrogen H-α light and (*b*) ionized calcium light. The plages (bright areas) and hydrogen dark filaments are clearly seen. (Sacramento Peak Observatory, Air Force Cambridge Research Laboratories)

9-11). One of the first signs is a local increase in brightness and the emission of short bursts of radio energy. After 15 to 30 minutes, the explosive phase begins, and for 15 minutes or more the flare emits short-wavelength X-rays, ultraviolet radiation, visible light, and radio energy in the meter, decameter, and centimeter range. Gases are ejected from the region at velocities in excess of 600 miles per second. Shock waves are propagated over the solar surface at velocities of 1,000 miles per second. Vast amounts of energy (about 10^{33} ergs) are emitted during a flare. Clearly, the Earth does not escape the effects of such an explosion.

The gases, composed mostly of ionized hydrogen, that is, protons and electrons, reach the Earth in a very short time. Some of the most energetic particles travel at nearly the speed of light and reach the Earth in just over 8 minutes. Other less energetic particles take longer. Some of the very energetic protons reach the Earth in an


2150:20 2204:10 2207:10 2211:20

2214:10 2227:00 2232:40 2325:50

Figure 9-11 The eruption of a solar flare as seen in hydrogen light; total time lapse is about 1 hour 35 minutes. (Lockheed Aircraft Corporation)

hour or so, indicating that they do not travel in straight lines. The path of these particles is presumably deviated by the sun's magnetic field. Upon arriving at the Earth, the low-energy particles are deviated by the Earth's magnetic field and enter the atmosphere in the polar regions. These particles are related in a manner not well understood with the *aurora*—the northern lights and the southern lights. The more energetic particles come right on through the magnetic field and strike the surface of the Earth. They have been known to play havoc with electrical equipment, such as causing circuit breakers for towns and cities to go out.

Ultraviolet and X-ray radiation disrupt the ionized layers in the Earth's atmosphere, which by reflecting radio broadcasting waves ordinarily enable radio operators in New York to receive messages from London, Sydney, Caracas. The ionized layers themselves are a result of the sun's usual ultraviolet radiation, but the intense radiation associated with solar flares has a disrupting effect and causes a corresponding fading in long-distance radio communications. Since we rely so heavily on long-distance radio communication (especially for ships and airplanes), there is a very practical purpose in observing the sun continuously: we may predict disruptions in the radio service and permit advance arrangement with other means of communication such as a communication satellite.

However, not all solar flares result in the corresponding terrestrial activity; nor does such terrestrial activity occur only when visible markings are in evidence on the photosphere. Magnetic disturbances often occur at intervals of 27 days, some of them when no corresponding visible markings appear on the sun. Since this 27-day interval between some of the magnetic storms is too frequent to be accounted for by mere chance, and since it is the synodic rotation period of the sun at 10° solar latitude, it is suspected that some regions of the sun remain disturbed for months at a time, even though they may not be visible.

The sun's photosphere, then, is mottled by granulation, marked with spots, and occasionally blemished by a flare. But even aside from these irregularities the photosphere does not present a uniformly bright disk when photographed. Its brightness decreases from the center of the disk to the limb. This is called *limb darkening* (Figure 9-9*b*). We cannot see into the sun as deeply at the limb as at the center of the disk, for at the limb light must travel through a thicker layer of gas. Consequently the visible regions of the limb are nearly 190 miles higher in altitude and consequently cooler than the center of the disk. It is this lower temperature that accounts for the decrease in brightness. By taking successive spectrograms of small regions of the sun from the center of the disk out to the limb, we actually obtain spectrograms of ever-higher levels above the photosphere and therefore learn more about the distribution of temperature and pressure on the sun.

Chromospheric Activity

We have seen that above the photosphere lies the chromosphere extending some 5,000 to 10,000 miles. The thickness of the chromosphere varies somewhat with the sunspot cycle. During a sunspot minimum its thickness at the poles is about 10% greater than at the equator, and during the maximum the thickness is the same all around the sun.

The fact that the bottom of the chromosphere is cooler than the photosphere does not establish a general tendency to decreasing temperature with increasing altitude. At an altitude of about 2,500 miles the temperature of the chromosphere is about 6,000°K. At its rather indefinite top where it merges with the lower corona, its temperature is closer to 50,000°K.

Like the layers immediately beneath it, the chromosphere is turbulent. When seen during a total solar eclipse or with a coronagraph (to be discussed) it presents a very uneven outline that has many small (in solar terms) irregularities called *spicules* projecting upward (Figure 9-12). These are several hundred miles in diameter and go as high as 10,000 miles. Their lifetime varies from 2 to about 12 minutes, with an average of about 3 to 5 minutes. Thus they have about the same diameter and lifetime as a bright granule. That the spicules are somehow related to the granules on the photosphere is difficult to prove but tempting to assume.

It is quite obvious that gases do extend up from the chromosphere (Figure 9-13). These projections, called *prominences*, were first seen during a total solar eclipse. Since prominences are such an obvious sign of solar activity they are very important in the study of the sun, but to photograph them only when the sun is shielded by the moon not only demands great patience on the part of the astronomer but used to mean that he could not study them systematically.

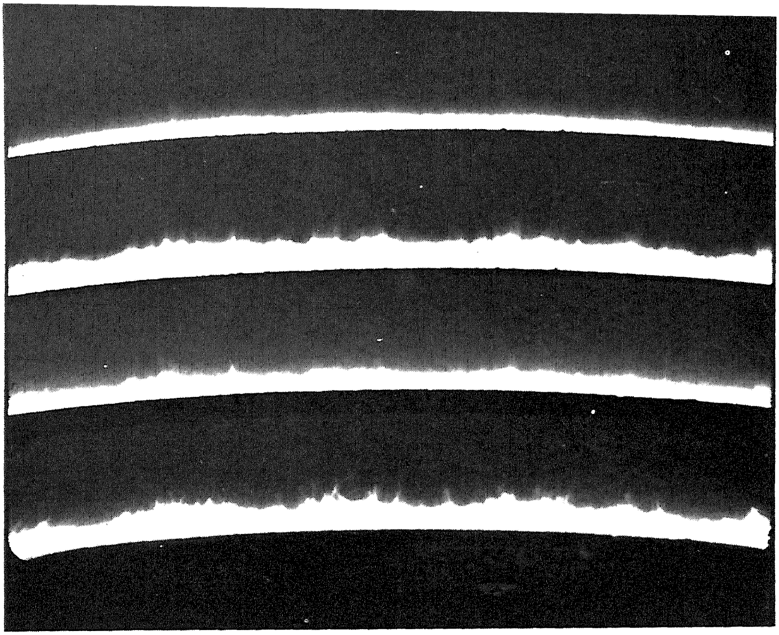


Figure 9-12 Spicules photographed with a 15-in. coronagraph. (Sacramento Peak Observatory)

Several types of instruments have been built to alleviate this difficulty. One of these, the *coronagraph,* was devised to produce an artificial eclipse during which the prominences and the inner part of the corona can be studied systematically. Since the sun's light is so bright that reflections from dust and air cause blurring, coronagraphs must be installed and used at very high altitudes where the air is thin and clear. The most famous have been placed in the Rocky Mountains at the High Altitude Observatory at Climax, Colorado, on the Pic du Midi high in the French Pyrénées, and on Sacramento Peak near Sunspot, New Mexico.

Since the coronagraph is able to simulate a solar eclipse whenever the sun is visible, it is possible to take time-lapse movies in which each frame is taken at a specified time interval (for example, 1 per minute). When such a movie is run through a motion picture projector at 16 frames each second the motions of the prominences are speeded

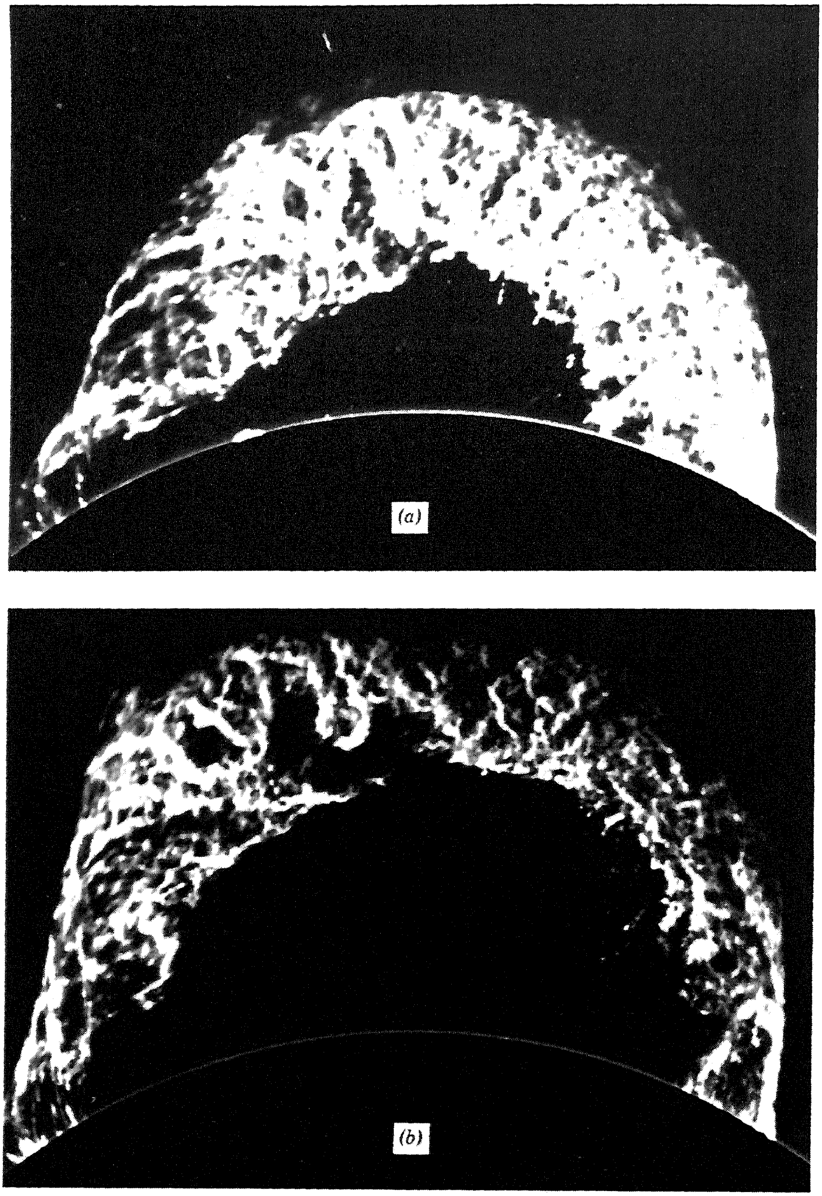


Figure 9-13 The great arch prominence of June 4, 1946. The two exposures were taken only 15 minutes apart. The diameter of the sun, roughly 864,000 miles, gives some idea of how rapidly this arch left the sun. (Sacramento Peak Observatory)

up. In this way it is possible to determine not only how fast the gas in the prominences moves but also along what path it moves.

By this means we find that there are many varieties of prominences: some look like geysers, some take the shape of loops, some appear as arches, and still others hang above the sun quiescently like stringy clouds.

The *arch prominence* is apparently a dark filament seen from the side. The most spectacular arch prominence yet photographed occurred on June 4, 1946 (Figure 9-13). The first of the time-lapse photographs was taken shortly after sunrise when the prominence had already risen to a height of 250,000 miles. The prominence moved upward from the sun at velocities in excess of 100 miles per second, so that it was a matter of only ½ hour before the prominence had extended beyond the limits of the coronagraph. This prominence was actually seen as a dark filament during the few days before it erupted into this fantastic upheaval (Figure 9-14).

A *surge prominence* (Figure 9-15*a*) may burst forth from the photosphere like a huge column of gas, stand with its top perhaps 30,000 miles above the photosphere, then fade and fall back into the sun along the same path. Such surges apparently result from some sort of explosion under the photosphere.

Loop prominences (Figure 9-15*b*) are associated with sunspots and are peculiar in that the material of both branches of the loop flows downward. Thus, gases seem to enter the loop at the top, divide, and flow into the associated sunspot along the two halves of the loop. Ionized gases falling into the sun are apparently constrained into the loop by the magnetic field about the sunspot.

Falling gases may also take the form of *coronal rain,* a flow of gases into the sun along definite paths that do not appear as loops. It is possible, however, that coronal rain may sometimes be loops seen edgewise. Often coronal rain will flow from different directions into a common region of the sun, as if the sun were "sucking" the gases in like a vacuum cleaner.

A *quiescent prominence* (Figure 9-15*c*) is also a dark filament, seen from the side. It may remain for months before it breaks up by either vanishing or exploding, like the great arch prominence of June 4, 1946. It seems rather amazing that on the turbulent sun quiescent prominences can remain so long with so little apparent change, and in fact, although their form may not vary, time-lapse

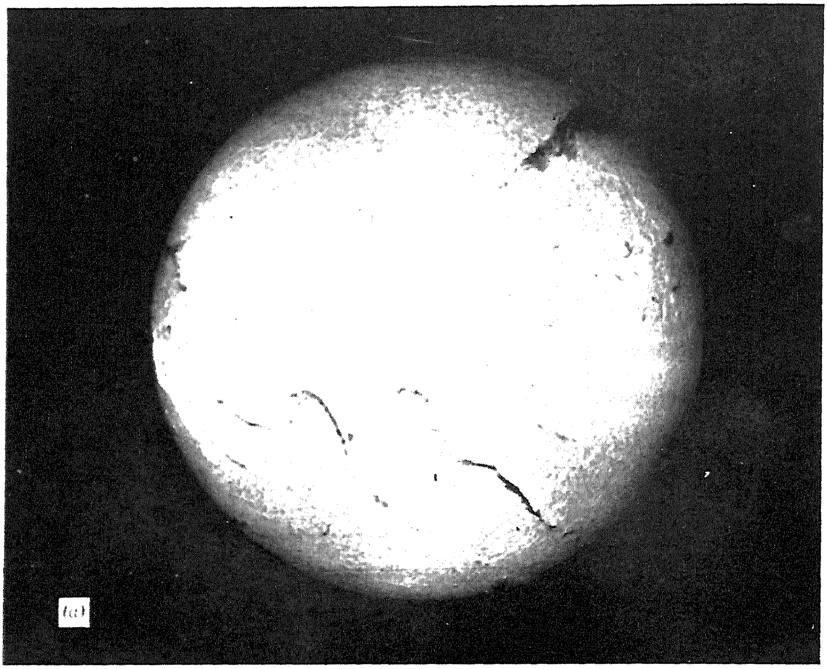
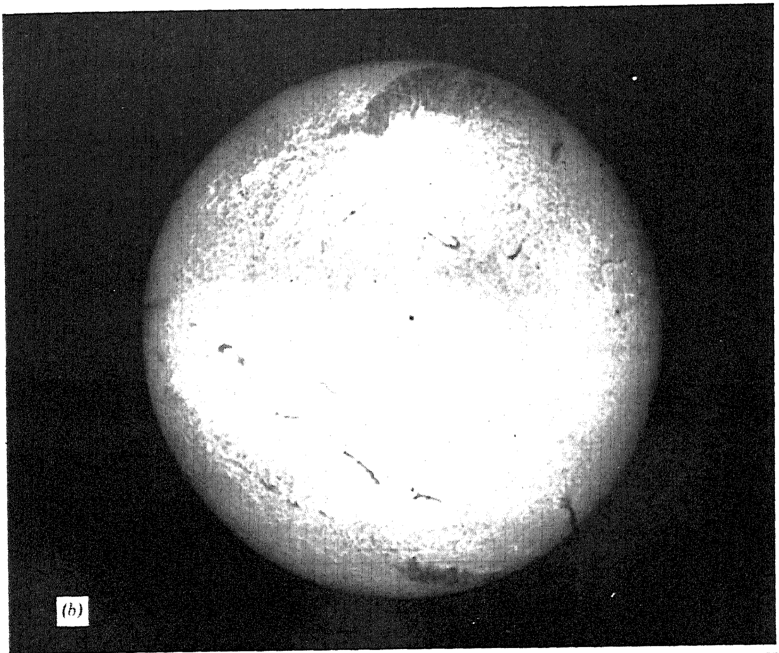
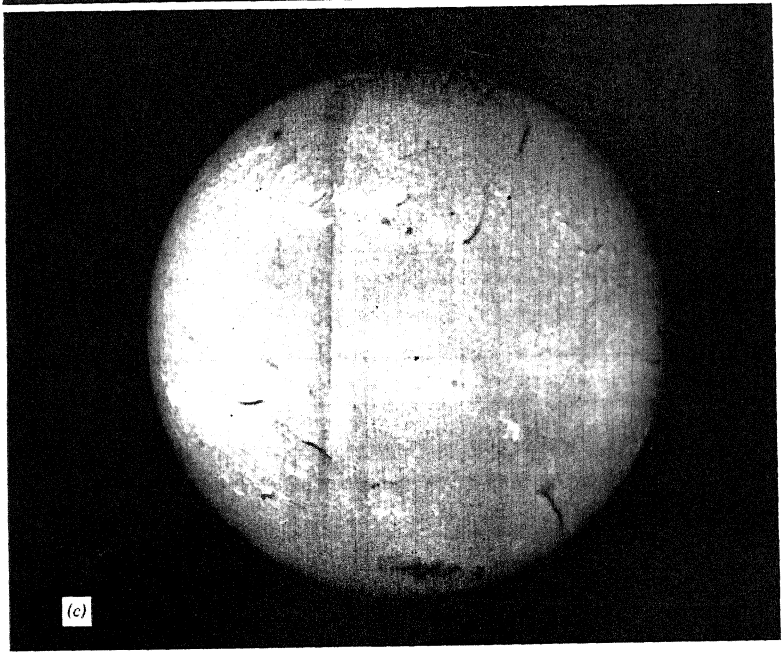


Figure 9-14 The great arch prominence was seen as a dark hydrogen filament on (*a*) May 30, (*b*) June 2, and (*c*) June 3, 1946. (Observatoire de Paris, Meudon)

movies reveal that the material composing them continually changes. Here, as in loop prominences and coronal rain, we find that gases reveal their presence by falling into the sun, becoming incandescent only when traveling in the region defined by the prominences.

It is not known precisely why gases should emit light only in a restricted region, but there is reason for thinking that this region is one whose temperature is lower than that of the inner corona which surrounds it. It is possible for a gas to be so hot that it becomes completely ionized (this occurs most readily with hydrogen). Hence, its atoms retain no electrons that may make transitions, and consequently the gas cannot emit visible light! When a gas is too cold, therefore, it does not emit visible light because there is not enough energy to excite the electrons, and when it is too hot it does not emit visible light because the electrons have become so excited that they have escaped the atoms completely. Between these two extremes the gas emits visible light. We have seen that the corona is hot enough


(b)


(c)

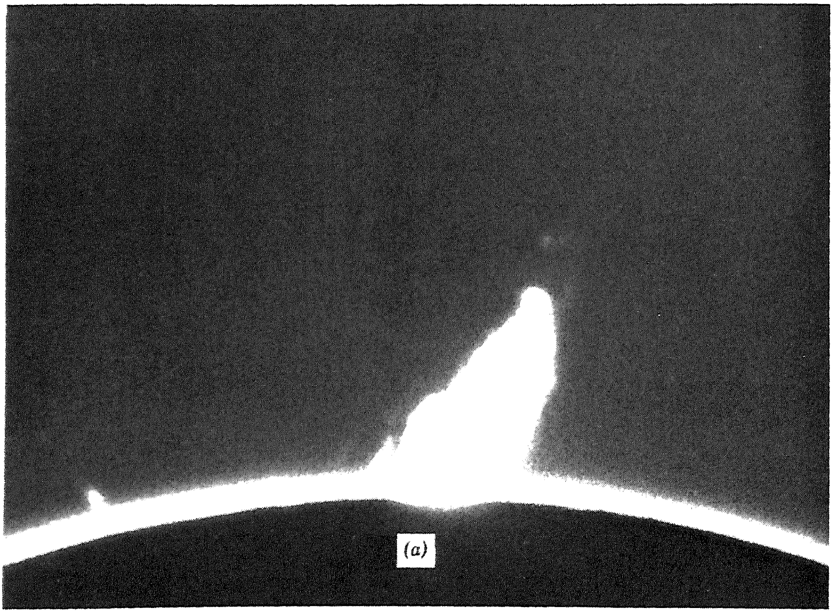


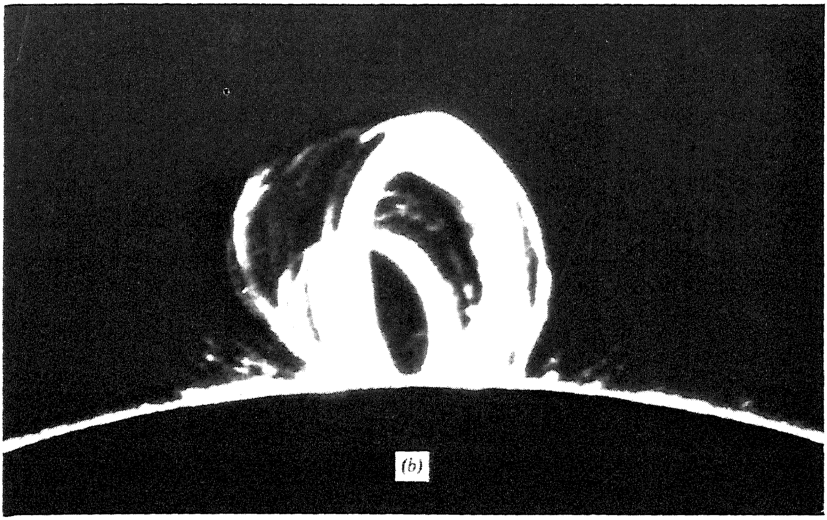
Figure 9-15 (*a*) Surge prominence. (High Altitude Observatory of the University of Colorado) (*b*) Loop prominence. (Sacramento Peak Observatory) (*c*) Quiescent prominence. (Sacramento Peak Observatory)

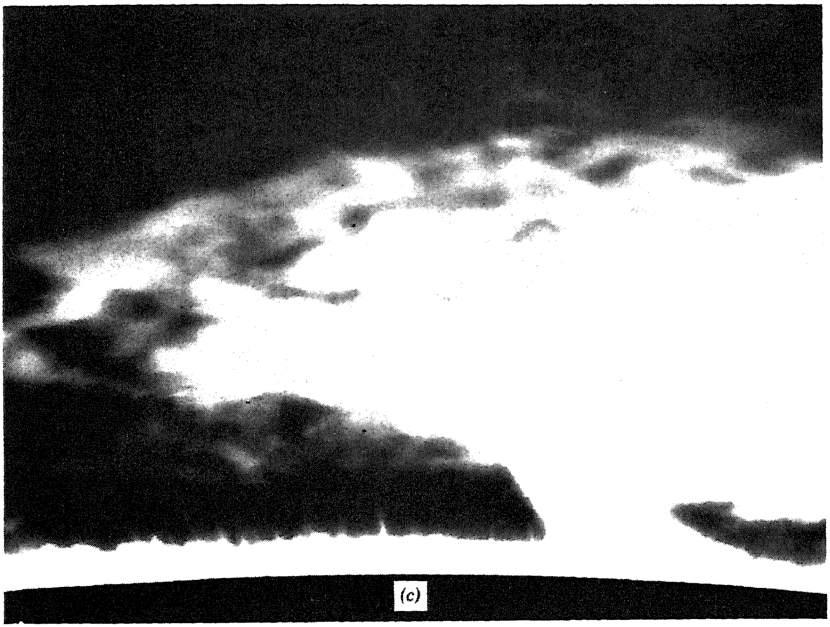
to ionize iron thirteen times; so it is hot enough to completely ionize hydrogen, the main constituent of the prominences. For this reason we suppose that gases falling through the corona must fall into a cooler region before they can emit visible light. Why these regions are cooler remains to be determined, but the study of magnetic fields may lead us to the answer.

The fact that gases *fall into* the sun to produce many of the prominences may seem puzzling. But we see only the rains on the Earth; we do not *see* water evaporating from the oceans and lakes. Perhaps gases, too hot to be seen, leave the turbulent photosphere in a like manner to go thousands or millions of miles above the sun and then fall back down to produce the "rain."

The Corona

Outward from the chromosphere there extends the pearly-white tenuous gas called the corona, which is seen so beautifully during


(b)


(c)

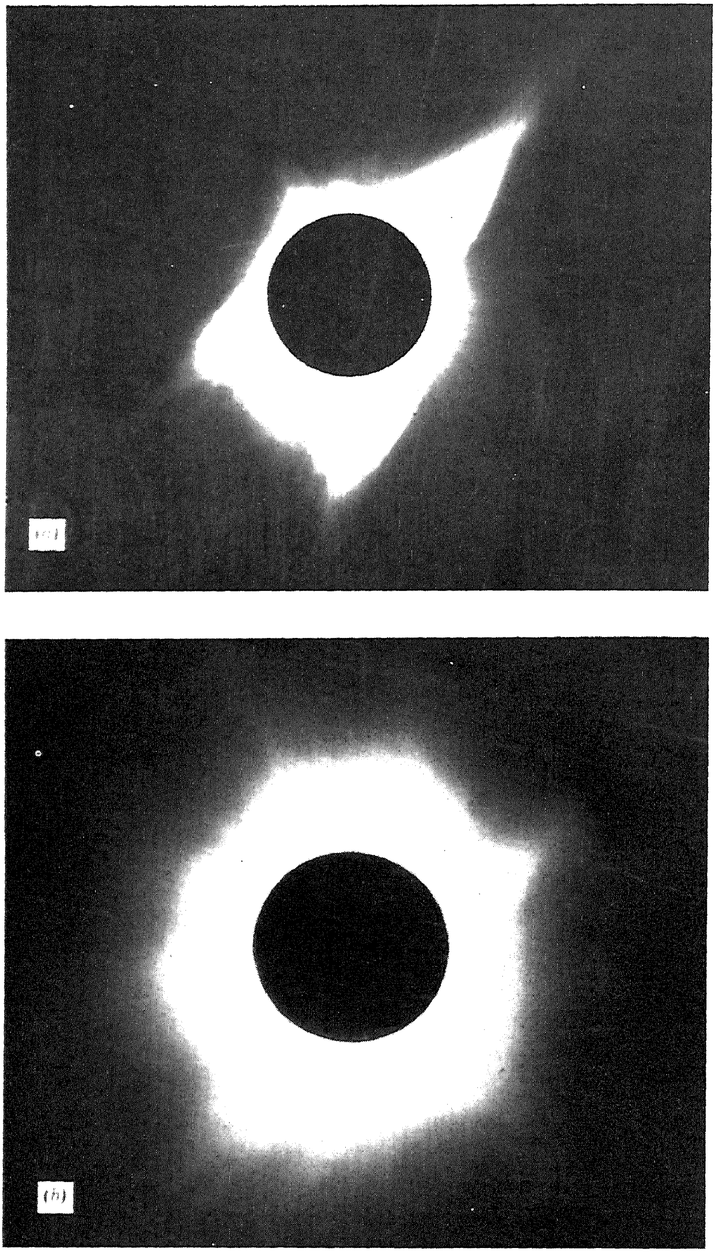


Figure 9-16 Total eclipse of the sun (*a*) during a sunspot minimum on February 25, 1952 (Yerkes Observatory); (*b*) during a sunspot maximum, October 12, 1958 (Courtesy of F. L. de Romaña, Peru)

a solar eclipse. Its height above the photosphere, as well as its shape, varies with the sunspot cycle. During a sunspot maximum it is fairly regular, but during a minimum it extends out much farther from the equatorial regions in huge petal-like configurations (Figure 9-16). The inner part of the corona has a brushlike appearance which seems shaped by magnetic lines of force near the polar region. Since the corona extends millions of miles into space, it may be associated with the particles (discussed on p. 180) which cause the zodiacal light.

The coronal spectrum consists of three components. The inner corona emits a pure spectral continuum with no absorption lines whatever, whereas the spectrum of the outer corona exhibits both bright lines and a continuous spectrum with Fraunhofer absorption lines.

Since the continuous spectrum from the outer corona is marked by Fraunhofer absorption lines, it must be the solar spectrum scattered from minute particles. The absorption lines, however, are broad and rather washed out, indicating that these particles are traveling with such a high velocity that they impart a considerable Doppler shift to the light that they reflect. Since the particles move in all directions, the lines are broadened simultaneously to the red and to the blue.

It is the emission spectrum that contains the lines of iron ionized thirteen times. We have noted (p. 208) that such ionization in the corona indicates not only a very high temperature of perhaps 1,000,000°K, but also a low pressure as well. The emission lines, like the absorption lines, are broadened by the rapid motion of the atoms that compose the very hot gases of the corona. This type of broadening is called *Doppler broadening* and is slightly different from turbulent broadening, in that the motions involved are not those of gaseous masses, as with turbulence, but are those of randomly moving particles.

Whether the pure continuous spectrum from the inner corona is caused by Doppler broadening so extreme that the lines become completely washed out, or whether it is caused by some form of electron capture is a problem that needs further study before a definite answer is found.

In order for the corona to maintain a temperature of roughly 1,000,000°K it must receive energy from the sun. Since it is quite transparent to electromagnetic radiations, it does not seem likely that it can be heated by this mechanism. It has been suggested that the turbulence of the photosphere sets up shock waves which are transported into the corona. The pressure of the corona is so low, however, that shock waves cannot continue as such; they die out and the energy they contain is transformed into heat energy.

The fact that the temperature of the corona increases with increasing altitude has led to some rather startling observations. An ionized gas at low pressure is a good conductor of heat. The conduction of heat outward from the sun results in the expansion of the outer corona. At about 6,000,000 miles from the solar surface, the corona is expanding at a speed in excess of 100 miles per second—faster than the speed of sound in the same gas. The outer corona is turned into a supersonic wind; the outer corona becomes the *solar wind* which whizzes past the Earth at a velocity roughly 250 miles per second. The solar wind continues out beyond the orbit of Jupiter, but just how far remains to be seen by future research. The existence of the solar wind has been verified by measurements made by spacecraft on their way to the moon, Venus, and Mars. The solar wind at the Earth's distance from the sun is composed of 10 protons per square cm, or perhaps more. A comet is truly a "solar wind-sock."

The sun has been and is now being studied by all available methods and with all the imagination and ingenuity that can be brought to bear on it. Very important to us here on the Earth is its 11-year cycle of activity with its sunspots and flares, prominence activity, varying coronal shape and size, and their corresponding effect on our planet. It has been found, for example, that the rate of growth of trees, which depends on the weather, varies in an 11-year cycle which corresponds to the cycle of the sun. In fact, a study and understanding of this growth as seen in tree rings has been used to date pieces of wood found by archeologists in their diggings.

The effect of solar activity on radio broadcasting reception has already been mentioned. Its effect on the weather of the Earth has led meteorologists, too, to become interested not only in the 11-year cycle but also in solar flares; there is reason to believe that those powerful eruptions affect our weather to an as yet undetermined degree. Furthermore, since large flares emit cosmic rays in such deadly amounts, they will also concern the space traveler who may find it healthier to venture out beyond our shielding atmosphere only during a sunspot minimum, when solar flares are quite rare.

The astronomer, however, studies the sun primarily because it is a star, the only one whose surface we can see and study in some detail. And stars, after all, are the astronomer's first calling, for they are the building stones of the universe.

Basic Vocabulary for Subsequent Reading

- **Chromosphere**
- **Convective envelope**
- **Corona**
- **Dark filament**
- **Flash spectrum**
- **Granules**
- **Limb darkening**
- **Photosphere**
- **Plage**
- **Prominence**
- **Solar flare**
- **Solar wind**
- **Spectroheliogram**
- **Spicules**
- **Sunspot**

For Further Reading

Abetti, Giorgio, *The Sun*, Crowell Collier and Macmillan, New York, 1957.

Bray, R. J. and R. E. Loughheed, *Sunspots*, John Wiley and Sons, New York, 1965.

Ellison, M. A., *The Sun and Its Influence*, Routledge and Kegan Paul, London, 1959.

Menzel, Donald H., *Our Sun*, Harvard University Press, Cambridge, Mass., 1959.

Struve, Otto, and V. Zebergs, *Astronomy in the 20th Century*, Chapter VII, Crowell Collier and Macmillan, New York, 1962.

Babcock, Horace W., "The Magnetism of the Sun," *Scientific American*, p. 52, February 1960.

Chapman, Sidney, "The Earth in the Sun's Atmosphere," *Scientific American*, p. 64, October 1959.

Evans, John W., "Motion in the Solar Atmosphere," *Sky and Telescope*, p. 321, June 1963.

Malitson, Harriet H., "The Solar Energy Spectrum," *Sky and Telescope*, p. 162, March, 1965.

Minnaert, M., "Solar Symposium at Sydney," *Sky and Telescope*, p. 151, March 1964.

Ney, E. P., "Balloon Observations During July Eclipse," *Sky and Telescope*, p. 251, November 1963.

Parker, E. N., "The Solar Wind," *Scientific American*, p. 66, April 1964.

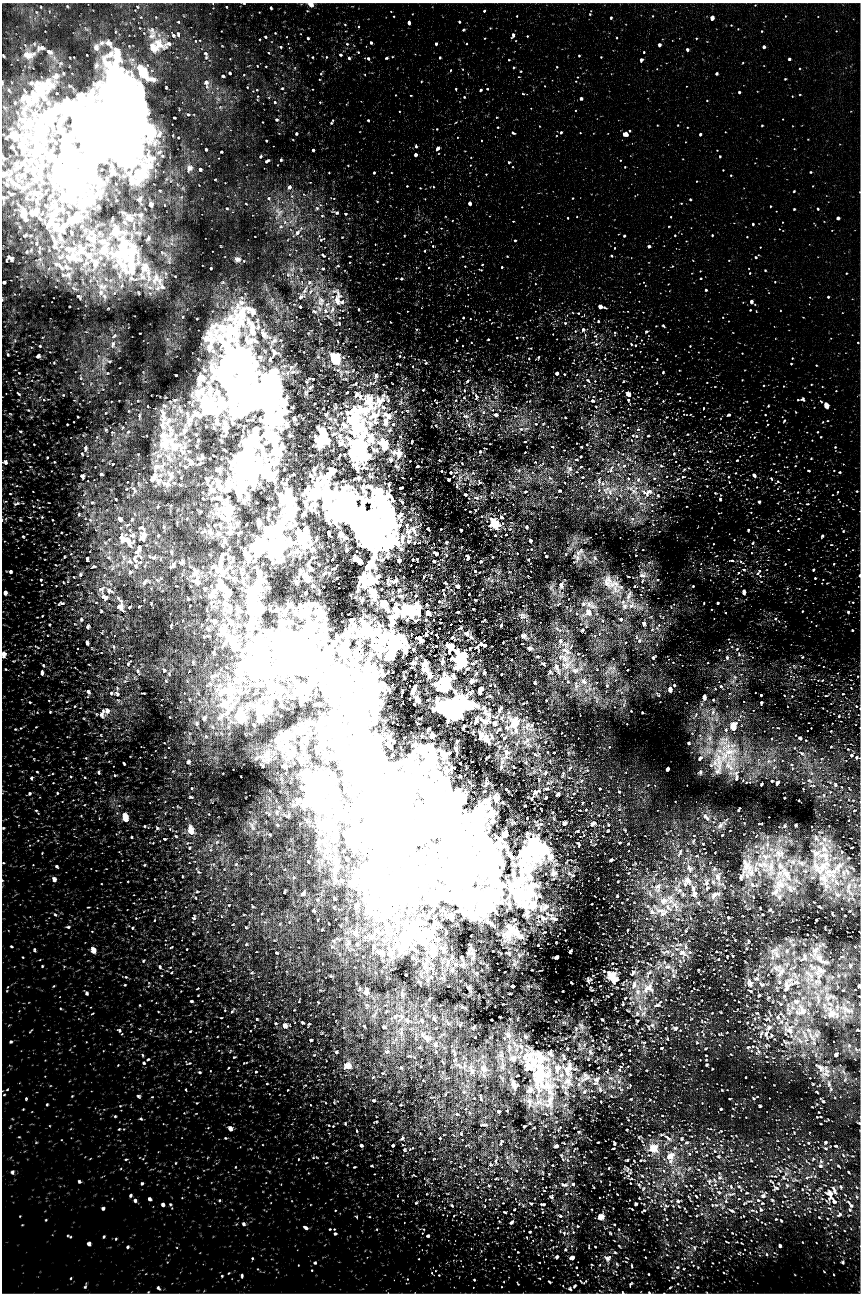
Schwarzschild, Martin and Barbara, "Balloon Astronomy," *Scientific American*, p. 52, May 1959.

Sirin, Harold, "Hot Spots in the Atmosphere of the Sun," *Scientific American*, p. 34, August 1958.

Warwick, C. S., "Solar Particles in Interplanetary Space," *Sky and Telescope*, p. 133, September 1962.

"Radar Observations of the Sun," *Sky and Telescope,* p. 70, February 1962.

"Solar Flare Symposium," *Sky and Telescope*, p. 89, February 1965.



Star clouds in the Sagittarius region of the Milky Way. The dark lanes are intervening gases that obscure the stars behind. (Mount Wilson and Palomar Observatories)

Chapter 10 *The Stars, a General Description*

Studying the stars is like putting together a jigsaw puzzle for which we do not have all the pieces. What is more, the pieces that are available to us do not simply lie around waiting to be picked up and fitted into place; they are often hard to find and when found their shapes are sometimes indistinct. The pieces at hand are the observed physical characteristics of the stars. When fitted together they leave some gaps in our picture. We then deduce what ought to fill those gaps so that what we see may make sense. The astronomer hopes thus to obtain knowledge of the stars' present state, of their origin and past, and of their future. He wants to understand how stars are born, what stages they undergo, and how they end.

The prodigious amount of energy that a star emits each second is evidence that it is not a dead and static body, but that it undergoes an evolutionary process. The basic assumption of science, that of causality, tells us that any emission of energy from a star must have a source which is consumed during the star's lifetime. As this source of energy is consumed the star must change.

In order to understand and predict stellar changes we must be able to describe a star. Such a description should include a knowledge of each of the following:

1. Distance from the sun or, more exactly, the position of the star in our galaxy.
2. Space velocity
3. Size
4. Temperature
5. Luminosity
6. Mass
7. Density
8. Chemical composition
9. Age.

Stellar Distances

The distance of a star from the sun is not only one of the more difficult characteristics to determine but also one of the more important. All the life processes of the star are determined by means of the amount and types of energy radiated, but the amount of energy a star radiates out into space cannot be known until its distance is known. Consequently, there has been a great deal of work done on the distance determination of stars, some of which has proved successful and some of which at one time was thought valid but was later proved incorrect. Hence, there is still much thought and work applied to this problem.

The first method used to determine a star's distance from the sun is the most fundamental, for it is an actual measurement of distance after the manner of the surveyor. It is the method of heliocentric parallax discussed on p. 104. If the nearer stars appear to oscillate back and forth in front of the background stars as the result of the Earth's revolution about the sun, we need only measure the *angle of parallax* p (Figure 10-1) through which a star appears to move. This angle when considered trigonometrically with the radius of the Earth's orbit as a base line yields the distance oS.

The need for an astronomical unit of measure to fit stellar distances becomes apparent when we consider that the nearest star is something

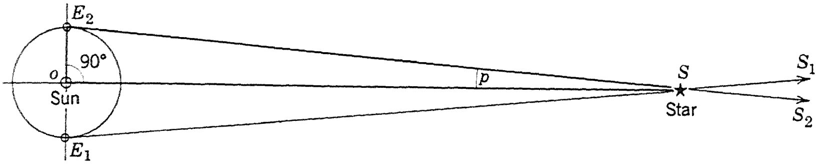


Figure 10-1 The angle of parallax is actually the angle subtended by the radius of the Earth's orbit as seen from a nearby star.

like 24 trillion (24×10^{12}) miles from the sun. The method of heliocentric parallax has given rise to a unit of distance called the *parsec.* One parsec is equal to about 19×10^{12} miles and is the distance at which a star must be in order to have a *par*allax of 1 *sec*ond of arc. The closest star must have the largest angle of parallax; it is called Proxima Centauri and its angle of parallax is 0.76 second of arc. Because all of the angles of parallax are very small, the distance d of the star in parsecs is defined as being equal to the reciprocal of the angle of parallax p measured in seconds of arc:

$$d = \frac{1}{p}$$

Proxima Centauri must be $1/0.76 = 1.3$ parsecs away. Another unit of measurement employed in stellar distances is the *light year,* or the distance light travels in 1 year. Light travels 186,000 miles every second or about 6×10^{12} miles in 1 year. A light year is, therefore, shorter than the parsec; 1 parsec equals about 3.26 light years. Proxima Centauri is about 4.2 light years from the sun.

As a method of determining distances, however, heliocentric parallax has inherent limitations. The farther the star is from the sun, the smaller the angle that must be measured. Since this angle must be measured by mechanical means (usually by the location of a star's image on a photographic plate) the measurement of the very small angles involved not only becomes difficult but the error in making the measurement soon becomes as large as the angle itself. Hence, only the distances of the nearer stars can be determined by heliocentric parallax. Distances for some 6,000 stars have been so determined, but only for the nearest 700 have they been determined with an uncertainty of less than 10%.

Astronomers also need to be able to determine distances for many of the billions of stars beyond the limit of heliocentric parallax. As we see in this and succeeding chapters there are many roundabout ways of determining stellar distances, all less accurate than the method of heliocentric parallax applied to the nearest stars. Yet each contributes its bit to our knowledge of the size of our galaxy and ultimately of the size of our universe. One of these methods (see Chapter 15) is based on the space velocity of stars, the next characteristic to be discussed.

Stellar Velocities

We have seen (Chapter 2) that the Doppler shift of the spectrum can be used to determine the velocity of a star in the line of sight, that is, its *radial velocity*. This gives the velocity in only one direction; but certainly stars do not only travel directly away from or toward the sun. An airplane, for example, that is flying northwest may be due south of an observer at one moment; shortly thereafter it will not only be closer to the observer but to the west of due south. The observer could say, therefore, that the plane has two simultaneous velocities: one to the north, approaching him, and the other due west. These two velocities are called *components* of the actual velocity. When we observe the radial velocity of a star, we observe only one of the components of its motion, that one equivalent to the airplane approaching the observer. The other component, the one equivalent to the westward motion of the airplane, is called *tangential velocity* and is seen as the star moves perpendicular to the line of sight. This motion is revealed by the star's motion against the background stars, which are too far away to reveal any tangential motion of their own. When tangential velocity is expressed as an angular motion per unit of time, for example, seconds of arc per year, it is called *proper motion.*

Figure 10-2 shows the radial velocity component, the tangential velocity component, and the relative velocity of a star with respect to the sun. Figure 10-3 shows the proper motion of a star that has moved against the background stars during the time interval between the three photographs. The spectrum of this star reveals its radial velocity. Its velocity relative to the sun is obtained by the Pythagorean theorem from its radial and tangential velocities.

The fact that stars do have proper motion interferes to a certain extent with the determination of their parallax, for in determining a star's position from year to year its proper motion must be untangled from its parallactic motion. Since parallactic motion is periodic and

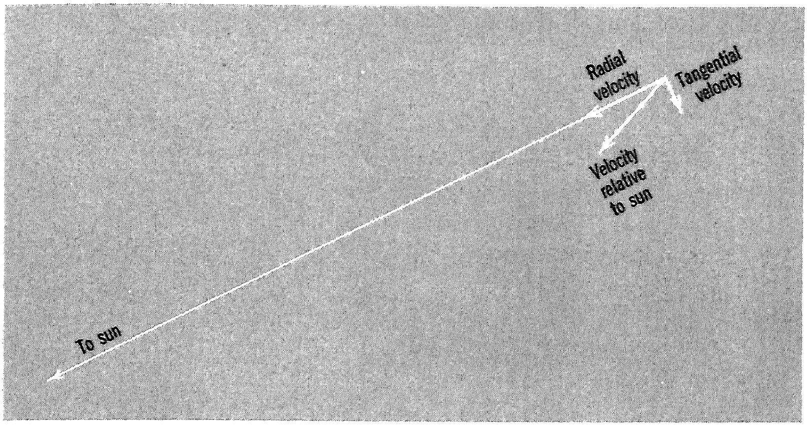


Figure 10-2 The velocity of a star with respect to the sun is broken up into two components—radial velocity and tangential velocity, for only these component velocities can be conveniently measured.

proper motion is cumulative, the separation of the two is not an impossible task. The proper motion of some 200,000 stars has already been determined.

Time is a factor in determining proper motion. The longer the time interval between photographs, the farther the foreground stars will have moved in front of the background stars. If one could wait 50 years between photographs, the proper motions of even quite distant stars could be determined by using the still more distant galaxies as a background against which a star's motion might be observed. Such a program is under way at the Lick Observatory. Beginning in 1947, the entire sky north of —23° declination was photographed over a period of 7 years on 1,246 plates, each plate covering about 25 square degrees of the sky. The original plan was to rephotograph the sky in 50 years, but with the advent of new and better equipment for measuring stellar positions, and because of some eagerness on the part of the astronomers involved, it was decided to rephotograph the sky after an interval of only 20 years. The results of this program should tell us a great deal about stellar motions. It should also enable us to learn a great deal about the rotation of our galaxy, for the more distant galaxies do not, of course, participate in this motion.

Relative velocity, as used here, refers to the velocity of a star with respect to the sun. If the sun's motion with respect to the local group

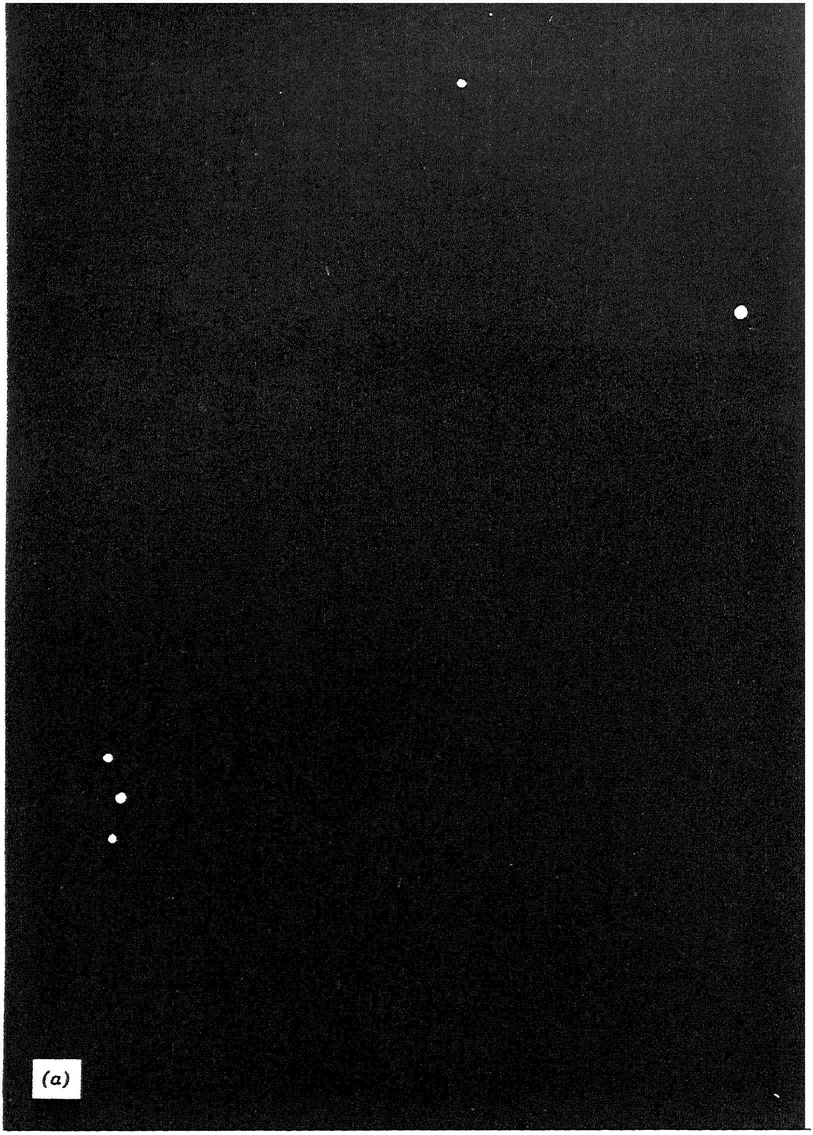
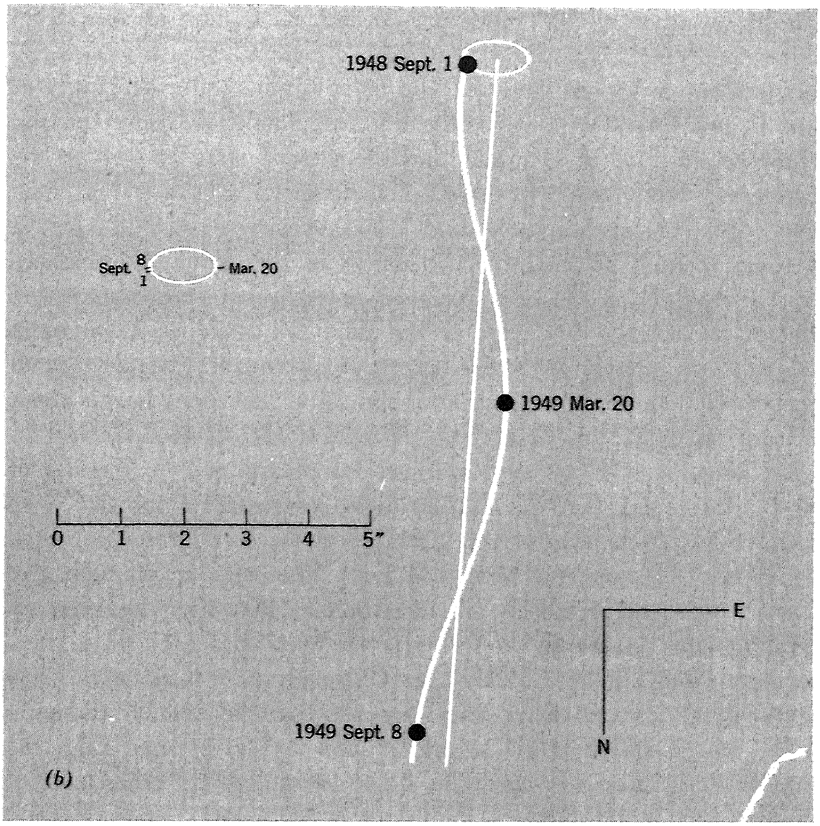


Figure 10-3 A composite of three photographs of Bernard's star (lower left). The three photographs have been superimposed so that the background stars (upper right) each show only one image. Motion to the north (down) is proper motion, motion east and west is the result of annual heliocentric parallax. (Courtesy Peter Van de Kamp, Sproul Observatory)



of stars is taken into account, that is, if it is subtracted from the relative velocity, we obtain the star's *peculiar velocity*—its velocity with respect to the local group of stars. The peculiar velocity of most stars in the neighborhood of the sun is about the same as that of the sun (about 12 miles per second) but there are certain stars, called *high velocity* stars, whose peculiar velocities are as great as 100 miles per second. As will be seen later, these are thought to be older than the sun itself. A star's velocity, as it revolves with the local group of stars about the center of our galaxy is called *rotational velocity*, since it is the star's motion as part of a rotating galaxy.

Stellar Sizes

The difficulties of studying parallax and proper motion are great because the stars are so very far apart in relation to their sizes and velocities. Although the velocity of a neighboring high-velocity star may be as much as 100 miles per second, its motion as seen from the Earth is very small indeed. Similarly, although some of the nearby stars are much larger than the sun, they appear only as points of light in our largest telescopes. To illustrate, let us imagine the sun shrunk to the size of an apple, and the nearest stars as well as the distances between each star shrunk in proportion. We would then have a group of apples, each about 1,500 miles apart, with the fastest of them moving at a rate of about ½ in. per hour. To measure such distances and velocities is hard enough, but it is even more difficult to measure the sizes of the stars; an apple at 1,500 miles does not look very large. Nevertheless the astronomer has set himself the task of determining the sizes of the stars; how does he go about it?

The angular diameter of a star can be determined by the use of an interferometer (see p. 40). In 1925, two smaller mirrors were mounted on 20-ft beam (it is called a *beam interferometer*) atop the 100-in. telescope on Mount Wilson. The angular diameters of 12 stars were measured by this instrument. Recently, however, an interferometer consisting of two separate telescopes each on a track has been installed at the Narrabri Observatory in Australia. Early reports from research in progress indicate that the angular diameter of the star Vega is 0.0037 second of arc! At its distance from the sun, its linear diameter must be 2.3×10^6 miles, 3.2 times that of the sun.

The diameters of some of the largest stars are astounding. Mira, in the constellation Ceti, has a diameter 460 times that of the sun. Were Mira to replace the sun, its boundaries would include the orbit of the planet Mars! The star Betelgeuse is a pulsating star and has a diameter which varies from 700 to 1,000 times that of the sun! The star α Hercules has a diameter 800 times that of the sun.

The diameters of smaller stars cannot be measured with interferometers. They are simply too far away and too small. In some instances, however, double stars revolve about their common center of mass in such a way as to eclipse one another as seen from the Earth. If the motions of the stars are well known, their diameters can be determined by the length of the eclipse (see p. 340). A star's diameter can also be estimated by the Stefan-Boltzmann law if its temperature and luminosity are known. The smallest stars have diameters less than that of the Earth, diameters roughly $1/500$ times that of the sun. Consequently, the range in diameters is rather large; the biggest stars

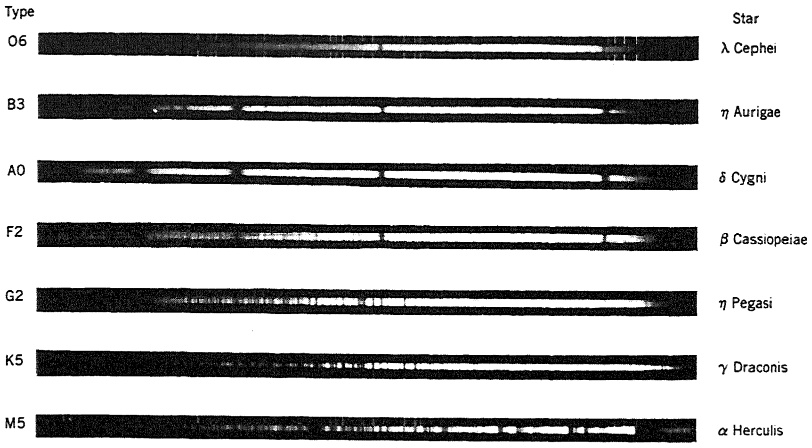


Figure 10-4 Principal types in the spectral sequence. (Mount Wilson and Palomar Observatories)

have diameters about 400,000 times the diameters of the smallest stars.

Stellar Temperatures

Stellar temperatures are fundamental in the study of the stars, for when a star's temperature is considered along with its size, we can find how much energy it radiates off into space and hence, for a stable star, how much energy is generated in its core. Fortunately, stellar temperatures are one of the more easily observable characteristics. The range of temperatures is less than that of diameters; the hottest known star, at about 100,000°K, is only about 70 times as hot as the coolest one (about 1500°K). Since the temperature of the sun is about 6,000°K, it may be classified as one of the cooler stars.

In Chapter 9 we saw that the temperature of a star determines its spectral characteristics. A hot star will have quite a different spectrum from that of a cool star. This can be clearly seen in Figure 10-4, where spectra are arranged according to the temperature of the emitting star. On the left hand side of the figure, each spectrum has a letter which is used to designate that particular type. An O-type star is the hottest and the M-type is the coolest.

The letters used to designate this *spectral sequence* do not really make any sense: O B A F G K M R N S. It would almost appear as though the astronomer responsible for such a sequence of letters either did not know the alphabet or was simply not concerned with order. Fortunately, neither is the case. The history behind the formation of this sequence of letters helps explain the matter; it also shows that astronomers, like anyone else, have a tendency to become bound by an established system of nomenclature.

When the study of stellar spectra received the impetus of photography it became feasible to try to arrange the different spectra in some kind of order. Since hydrogen lines are present in so many of the spectra is was felt that these lines might be used as a criterion. Accordingly, the stars with the strongest hydrogen lines were called A-type stars, those with slightly weaker hydrogen lines B-type stars, etc. At the time this scale was established the reason for the differences in stellar spectra was not well understood, and only later did it become apparent that they represented differences in stellar temperatures. Further study revealed that an A-type star is not the hottest; a B-type star is hotter, and an O-type star is the hottest of all. So the order of the spectra was rearranged according to temperature, although the letters designating the types were retained except for a few that had to be discarded. Thus, we see that the spectral sequence is basically a temperature sequence.

Since the spectral sequence is so important in astronomy a scheme has been devised to help the beginning student remember it. If we let the letters in the sequence, O B A F G K M R N S, be the first letters of each word in a well-chosen sentence we get the following: *Oh Be a Fine Girl Kiss Me Right Now, Smack!* It is a rare student who forgets this, after even a cursory glance, and therefore it has so far survived the vicissitudes of expression in the English language.

Actually the R-, N-, and S-type stars are a branch of the spectral sequence. These stars have chemical compositions slightly different from the rest and are relatively cool.

To explain the dependence of stellar spectra on temperature we can start with the coolest, the M-type star, and follow the changes as we move to the hotter stars. It should be recalled that in a gas of a given temperature the atoms are not all equally excited. Some atoms, because of collisions with other atoms, will have more energy, and some, for the same reason, will have less. The temperature is based on the *average* excitation of the atoms in a gas. This difference in excitation of the atoms in each star accounts for the fact that the spectral (temperature) sequence is a continuous one, free of any sharp breaks.

The M-type stars have a temperature of about 3,400°K, which is too low to excite electrons to very energetic levels. Thus the lines of the most easily excitable atoms (mostly those of the metals, which have many lines in their spectra) predominate the spectrum. Furthermore, the M-type stars are cool enough to allow some of the atoms in their chromospheres to combine and form molecules. The bands of these molecules, especially those of titanium oxide (TiO), are stronger in this type than in any of the other stars.

The K-type stars have a temperature of about 4,400°K. At this higher temperature some of the molecules have been broken up and therefore the molecular bands are not as strong as in the type M. Spectral lines resulting from more highly excited metallic atoms appear, and some metals become ionized at this temperature. In fact since calcium is so easily ionized, the H and K lines of ionized calcium are not only very intense in K-type spectra but are at their strongest here. The M-type star is not hot enough to ionize calcium to such an extent, whereas the hotter stars will begin to ionize calcium doubly or at least cause the outer remaining electron to make transitions that do not result in lines in the visible part of the spectrum.

The hydrogen lines of the K-type star, although weak, are stronger than in the M-type star. Since the first line of the Balmer series ($H\alpha$) is in the red region of the spectrum it may seem logical that it should appear strong in the cooler stars. It must be remembered, however, that this line (in absorption) is the result of a transition from the second to the third energy level and cannot appear until some of the hydrogen atoms have already been excited to the second energy level, which is a process requiring more energy than is generally available in the cooler stars.

The G-type stars have a temperature of about 5,500°K. As the temperature increases to this value, the hydrogen lines become more intense, for more of the hydrogen atoms are excited to the second energy level. The molecular bands essentially disappear except for bands of cyanogen (CN) and CH which, though weak, are still present. The H and K lines of calcium become slightly weaker than in the K-type star, since the outer remaining electron in the calcium atom is excited to levels higher than those that produce the H and K lines.

The F-type stars have a temperature of about 6,600°K. At this higher temperature the hydrogen lines increase in strength, the lines of most ionized metals are intensified, and the neutral metallic lines become weaker as more of the metallic atoms become ionized.

The A-type stars have a temperature of about 9,500°K. In this spectral type the hydrogen lines reach their maximum strength, and the lines of most of the singly ionized metals begin to weaken; the atoms become more highly ionized and produce lines in the far-ultraviolet region of the spectrum. The lines of neutral metals are quite absent from the A-type stars, since most of the metals have become ionized.

The B-type stars have a temperature of about 15,000°K, at which the hydrogen atoms start to become ionized and the lines of the Balmer series thus begin to decrease in strength. The lines of neutral helium appear, as well as those of ionized oxygen and carbon.

The O-type stars have a temperature so high that it is difficult to determine. The estimates run somewhere around 50,000°K or more. The lines that primarily distinguish the O-type spectrum are those of ionized helium, for they appear in no other spectral class.

The temperatures given for the various spectral types are really averages; each spectral class embraces a considerable range of temperatures. For example, the hottest A-type star has a temperature of about 11,000°K while the coolest is at about 8,000°K. This range in temperature gives rise to corresponding changes in the spectra. To account for these differences, each major spectral class is subdivided. There are stars, for example, that are classified as A0, others as A1, A2, A3, etc. An A3 star is cooler than an A2. Such subdivisions within each spectral type are recognized by differences in line intensities, the hydrogen lines being strongest in type A0.

Figure 10-5 shows how the line intensities of some of the more important lines or groups of lines vary throughout the spectral sequence. The lines of each element or group of elements reach a peak of intensity in some particular spectral type, that is, at a particular temperature. This temperature is related to the amount of energy needed to excite the atoms of the element in question to the level at which transitions producing these lines are most likely to occur. At this point, then, the lines will be at their maximum intensity. It should be noted once again that the intensity of lines from an ionized element or group increases as the lines for the neutral element or group weaken.

A peculiar terminology has arisen along with the spectral classes. Very often astronomers want to speak not of one particular spectral class but of a group of spectral classes or subclasses. It was once

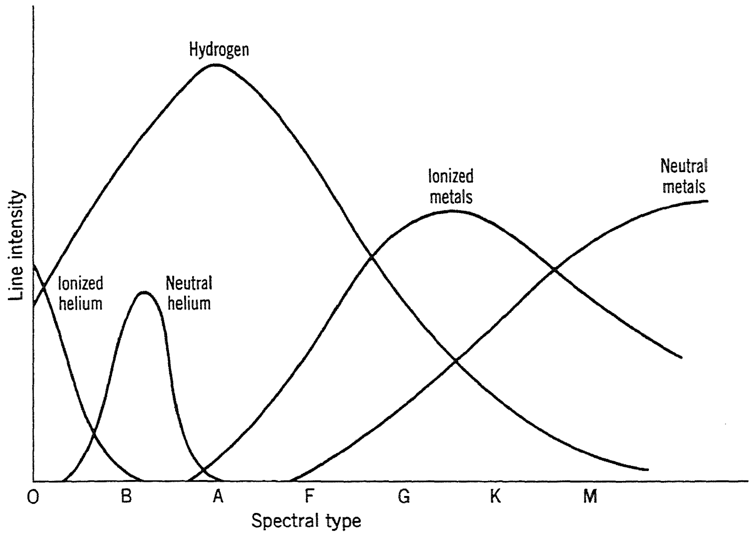


Figure 10-5 The intensity of a given stellar spectral line depends on the temperature of the star. Lines resulting from ionized atoms of a given chemical element will be more intense in a hotter star, and the lines of the neutral atoms will be more intense in a cooler star.

thought that the O-type stars were the youngest and that as stars grew older they cooled progressively into M-type stars. Therefore the stars classed as O, B, and A are loosely spoken of as *early-type* stars, and the K and M stars are referred to as *late-type* stars. An F0 or F1 star is sometimes called an early F star, but an F8 may be referred to as a late F star. This terminology is one of convenience only and has no strict definition or hidden meaning. It will be seen in Chapter 12 that an O or B star is indeed likely to be younger than a type M, although this need not be so.

Stars of very low temperature have recently been discovered and more are expected to be found as current researches continue at a number of observatories. These stars radiate mostly in the infrared region of the electromagnetic spectrum. Special equipment has been set up at Mount Wilson Observatory to survey the sky in search of *infrared stars*. Of those found so far, temperatures as low as 1,000°K have been measured. What sort of stars are these? Why are they so cool? These are questions being asked by astronomers.

Stellar Luminosities

The temperature of a star, according to the Stefan-Boltzmann law, determines how much energy is emitted per unit area of the star's surface. Of two stars the same size, the hotter will radiate more energy; of two stars the same temperature, the larger one will radiate more energy. The *luminosity* (intrinsic brightness) of a star, then, depends on two factors: its temperature and its size.

A star's *brightness*, however, is simply its apparent brightness as seen in the sky and depends not only on its luminosity but also on the distance between it and the sun. If we had three stars of the same luminosity at distances of 1, 2, and 3 parsecs, their brightnesses would progressively decrease with increasing distance according to the *inverse square law*. That is, if the brightness of the first star were 1 (on an arbitrary scale), the brightness of the second star would be $1/4$ which equals $(1/2)^2$, and that of the third star would be $1/9$ which equals $(1/3)^2$. If a star of the same luminosity as these three were 100 parsecs away, its brightness, on the same scale, would be $1/10{,}000 = (1/100)$.^{2} If the luminosity of any one of these were twice as great, its brightness would be twice as great.

The first astronomers concerned enough to make a record of the brightnesses of stars as they saw them were the Alexandrian Greeks. In the absence of a means for determining a star's distance from the sun they considered its brightness only. It was Hipparchus who first established a scale, called *apparent magnitude*, for comparing the brightnesses of stars. He did this in the course of compiling the first known star catalog, which contained the positions and brightnesses of just under 1,000 stars. The brightest stars he arbitrarily called of the *first magnitude;* the faintest stars visible to the naked eye were of the *sixth magnitude*, and all other stars were of intervening magnitudes according to their brightnesses. The system worked quite well, for it is fairly easy to estimate the brightness of a visible star on this scale. But with the advent of the telescope and of photography, the scale has had to be extended to ever fainter stars, until the *faintest* stars now photographed are of about the twenty-third magnitude.

Hipparchus misjudged the magnitudes of some of the brighter stars, however. When the magnitude scale was extended and expressed by a mathematical formula, it developed that the brighter stars are brighter than those of the first magnitude; indeed they are even brighter than those of zero magnitude. The only way to express these hitherto unsuspected magnitudes and yet retain the old scale (thereby avoiding a wholesale revision of stellar catalogs) was to adopt a few negative magnitudes. As a result, Sirius, the brightest star in the sky aside from the sun, has an apparent magnitude of —1.4 and the star 61 Cygni A has an apparent magnitude of +5.4 and is just visible to the naked eye on a clear moonless night when viewed from a place far from the lights, smoke, and smog of our modern cities.

The upshot is that a magnitude scale set up some 2,000 years ago is still in use, although stretched somewhat on a mathematical rack. After the revisions, a zero-magnitude star is exactly 100 times brighter than a fifth-magnitude star, which is fairly close to Hipparchus' scale. Thus, for a magnitude difference of 5 there is a change in brightness by a factor of 100. Consequently, a difference in magnitude of 1.0 corresponds to a factor in brightness of $(100)^{1/5} = 2.512$. A star with an apparent magnitude of 2.0, for example, is about 2.5 times brighter than a star with an apparent magnitude of 3.0. That star with an apparent magnitude of 2.0 is about $(2.5)^2 = 6.3$ times brighter than a star with an apparent magnitude of 4.0, about $(2.5)^3 = 16$ times brighter than a star with an apparent magnitude of 5.0 etc. A magnitude difference of 10 amounts to a difference in brightness of 100^2 or 10,000; a magnitude difference of 15 amounts to a difference in brightness of 100^3 or 1,000,000. Figure 10-6 shows a star field with indications of magnitudes for some of the stars.

There are stars over the entire sky which have been chosen as standard stars for magnitude determination. The magnitude of any star can be determined by comparing it with one of the standard stars which appear in the same part of the sky as the star being studied. The magnitudes are determined by comparing the electric currents set up when the star's light strikes a photoelectric cell. The brighter the star the greater the electric current produced in the cell. If a star whose apparent magnitude is to be determined produces an electric current equal to one of the standard stars, it has the same apparent magnitude as that standard star.

Before we can determine a star's luminosity (as distinguished from its brightness), its distance must be known. The scales used to measure luminosity are as arbitrary as the scale of apparent magnitude, and indeed one of the scales of luminosity, called *absolute magnitude,* is based on the scale of apparent magnitudes. It is an indication of the total amount of light radiated by the star and not the amount we receive here on the Earth. Unlike apparent magnitudes, absolute magnitudes are independent of the star's actual distance. By definition *the absolute magnitude of a star is the apparent magnitude the star would have, were the star 10 parsecs from the sun.* On this scale, then, all stars are compared from the same standard distance. If a

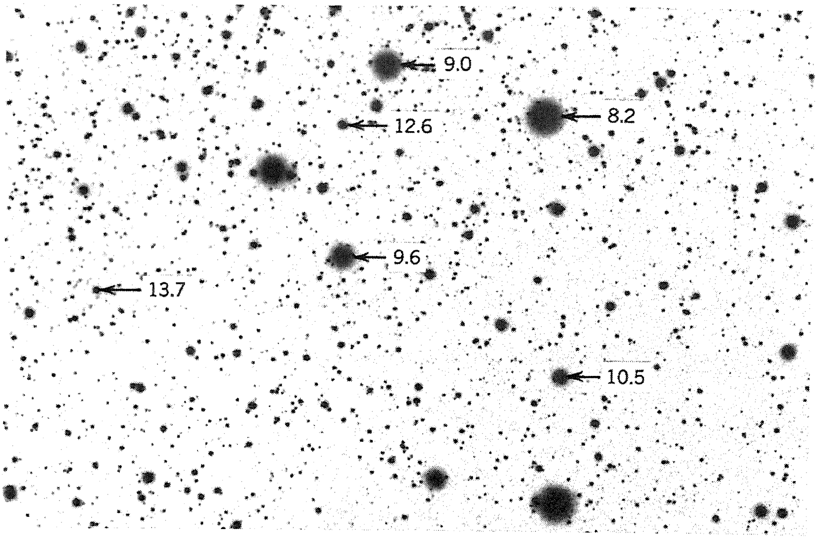


Figure 10-6 The same star field shown in Figure 2-12 with apparent magnitudes marked. The size and darkness (astronomers work with photographic negatives) of a star's image can be used to estimate the magnitude. (Courtesy of M. Walker, Lick Observatory)

star is closer than 10 parsecs, its apparent magnitude is numerically less (that is, it is brighter) than its absolute magnitude. If a star is more than 10 parsecs distant, its apparent magnitude is numerically greater (fainter) than its absolute magnitude.

Sirius, for example, is 2.7 parsecs from the sun. Its apparent magnitude is −1.4 and its absolute magnitude is +1.5. This means that if Sirius were moved away to 10 parsecs from the sun its apparent magnitude would be +1.5. Pollux, with an apparent magnitude of +1.1, is just barely more than 10 parsecs from the sun and thus its absolute magnitude is +1.0. On the other hand, Rigel, with an apparent magnitude of +0.14 is 200 parsecs from the sun so its absolute magnitude is −6.4. Since Rigel has an absolute magnitude of about 8 less than Sirius, it is about 1,600 times more luminous. The sun is so close that it has an apparent magnitude of −26.7, even though its absolute magnitude is only +4.9. Rigel's absolute magnitude is therefore 11.3 magnitudes less than that of the sun, that is, it is 33,000 times as luminous.

Paralleling the scale of absolute magnitude is another scale whose reference point is the sun, which is given an arbitrary luminosity of 1.00. A star with a luminosity of 10 would be 10 times more luminous than the sun; one with a luminosity of 1/100 is 1/100 as luminous as the sun.

The range in the luminosities of the stars is staggering. The most luminous star is about 10^{11} times more luminous than the least luminous star. The most luminous stars are about 100,000 times more luminous than the sun, whereas the least luminous stars known have a luminosity about 1/1,000,000 that of the sun. The sun, then, can be classified at about the middle of the range of luminosities.

For the study of *intrinsic* characteristics of the stars, luminosity (absolute magnitude) has a great deal of meaning, whereas brightness (apparent magnitude) has essentially no meaning because of the different distances involved.

The H-R Diagram

It is of interest to graph the stars by plotting their luminosities against their temperatures (Figure 10-7). Since a star's luminosity depends on both its temperature and size (see p. 253f), a graph plotting temperature against luminosity separates the stars according to size. If two stars have the same temperature but different diameters the smaller will be less luminous, since it has a smaller surface area and thus is located lower on the graph. Conversely, if two stars have the same luminosity but different temperatures the cooler must be the larger. Since it is cooler it is located farther to the right on the graph.

Such a graph is called an *H-R diagram* in honor of Hertzsprung and Russell, the two astronomers instrumental in compiling the first such diagram. Because of its tremendous significance in astronomy we shall refer to the H-R diagram frequently.

We notice that there is a great number of stars forming a sequence from the upper left to the lower right of the diagram. This band is called the *main sequence,* because it includes the majority of the stars observed. Above the main sequence we see a rather large group of stars classified as *giants* and *supergiants.* The supergiants, as the name indicates, are more luminous and thus larger than the giants, which are in turn larger than the main-sequence stars of the same temperature. Below and to the left of the main sequence, we see a group of stars that must be smaller than the main-sequence stars of the same temperature and thus are called *white dwarfs.* The late main-sequence stars (the sun included) are often called dwarfs because of their smallness when compared with giants of the same temperature. Nevertheless they are larger than the white dwarfs and

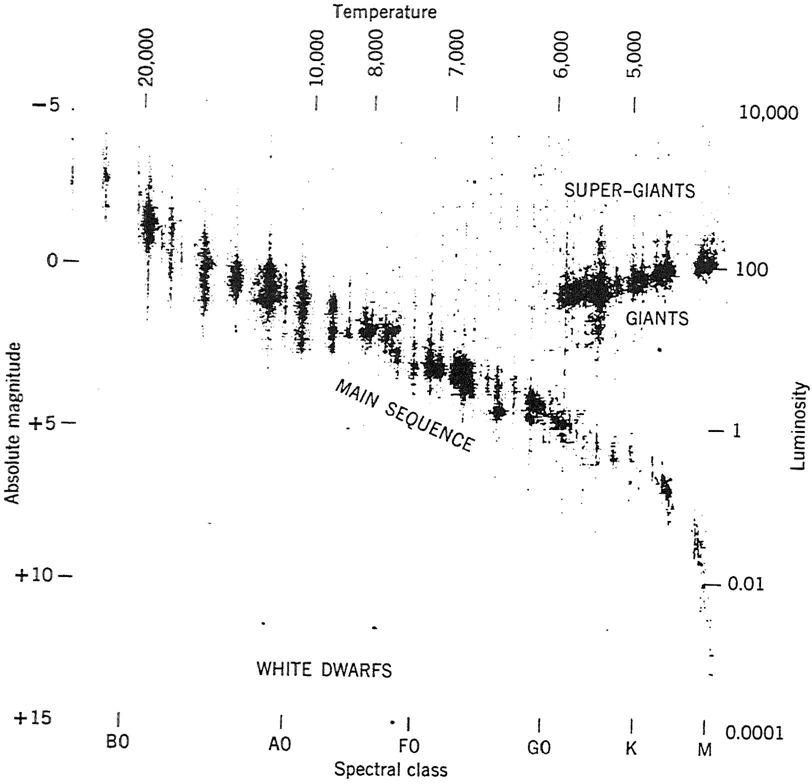


Figure 10-7 The H-R diagram. The temperatures refer to the main-sequence stars, and the grouping of points about vertical lines results from the method of spectral classification. (Courtesy of the late O. Struve from a diagram by Gyllenberg, Lund Observatory)

are not white in color.

That such a distinctive grouping of stars should exist when they are placed on a graph according to their temperatures and luminosities cannot be mere coincidence; there must be some meaning to this arrangement. The feeling at present among most astronomers is that these groupings represent the stages that a star goes through in its evolution from birth to death. We shall examine this process in Chapter 12.

Spectroscopic Parallax

One of the more beneficial results of the H-R diagram is that it can be employed to find the distances of stars too far away for heliocentric parallax determination. This is possible because it gives us the luminosity of a star, and once the luminosity is known we can find the star's distance by measuring its brightness; for brightness, luminosity, and distance are related very nicely by the inverse square law. By using luminosity units (as against the system of magnitudes) the brightness B of a star in the sky is proportional to the star's luminosity L and inversely proportional to the square of its distance D from the sun:

$$B \propto \frac{L}{D^2}$$

This proportionality can be converted into an equality by inserting a constant of proportionality K:

$$B = K\frac{L}{D^2}$$

The numerical value of K depends upon the choice of units for the brightness B and L. If L is based on the sun's luminosity, B can be based on a group of stars just as the apparent magnitude is.* This equation, expressed in units of luminosity and brightness permits us to see clearly how the brightness of a star depends upon its luminosity and distance. If both the brightness and luminosity of a star are known, its distance can be found by algebraic manipulation of the equation

$$D = \sqrt{\frac{K \cdot L}{B}}$$

But the problem of determining the luminosity of a star without first knowing its distance remains. To do this we must find some way of determining, independently of its luminosity, whether the star is a dwarf (late main-sequence), a giant, or a supergiant. If this can be determined, a star's luminosity can be found by correlating its spectral type with its size. For example, if a K0 star is estimated to be a dwarf, its luminosity must be about 0.4; if a K0 star is a giant, its luminosity will be a little less than 100 (see Figure 10-8). But how is it possible to determine the size of a star from a study of its spectrum?

It was pointed out that the diameters of the stars vary by a factor of more than 100,000, but, as we shall see later in this chapter, their masses vary by a factor of only about 1,000. Thus, the very large

---

* An equivalent expression used by astronomers which relates the apparent magnitude m of a star to its absolute magnitude M and its distance d (in parsecs), is logarithmic in nature: $M = m + 5 - 5 \log d$.

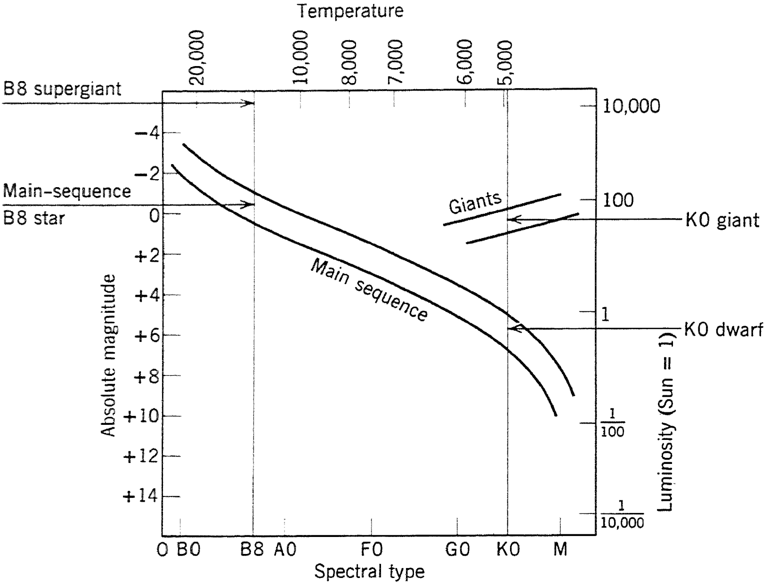


Figure 10-8 Spectroscopic parallax depends on the ability of the astronomer to determine from a star's spectrum alone whether a given star is a main-sequence star, a giant, or a supergiant. Once this is determined the H-R diagram yields the star's luminosity. The temperatures refer to the main sequence.

stars are large in size only and are not as massive as their sizes might suggest. They must, therefore, be expanded versions of the more numerous main-sequence stars. Because they are enlarged they must have smaller densities and thus lower pressures. It seems likely that this difference in pressure would affect the spectrum by producing different amounts of pressure broadening of the spectral lines (discussed on p. 56).

By placing the spectra of a main-sequence star and a supergiant of the same spectral type side by side, we do indeed find differences in the spectra (Figure 10-9). Not only are some of the lines of the main-sequence star's spectrum a little broader, but also some of the lines of ionized metals in the spectrum of the supergiant are more intense.

Both these results are to be expected. The main-sequence star has a higher pressure and some of its lines suffer some consequent pressure broadening. The supergiant has a lower pressure and more of its ions are able to remain ions; they and the electrons are farther apart, and thus it is more difficult for them to recombine and form neutral atoms. Since the intensity of ionized lines is one of the criteria

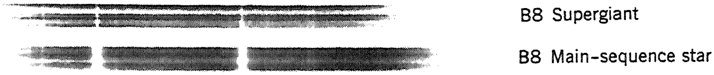


Figure 10-9 The spectral lines in a B8 main-sequence star are noticeably broader than those in a B8 supergiant. The most obvious lines in each of these spectra are those of hydrogen. (Yerkes Observatory)

for spectral classification and since ionization depends not only on the temperature of the star but also on the pressure, the temperature of a giant star is lower than that of a main-sequence star of the same spectral type. The giant star has more intense lines of the ionized atoms and is classified accordingly. A K0 dwarf, for example, has a surface temperature of 5,000°K; the temperature of a K0 giant, however, is 4,100°K, and of a K0 supergiant only 3,800°K.

It is therefore possible to distinguish between a dwarf, a giant, and a supergiant by an examination of their spectra. Once a spectrum has been classified into its spectral type and examined to determine whether it originates in a dwarf, a giant, or a supergiant, the star can then be placed on the H-R diagram. Once placed on the H-R diagram according to its spectral type and size, its luminosity can be read off directly.

The spectra shown in Figure 10-9 are both of B8 stars, one a main-sequence star and the other a supergiant. By referring to the H-R diagram (Figure 10-8) we see that the luminosity of a main-sequence B8 star is about 100. The supergiant, on the other hand, has a luminosity of about 30,000. Actually there is no sharp dividing line between the giants and the supergiants; the extent of pressure broadening and the extent of ionization can be used to classify the star with greater accuracy than merely placing it in one of only two classifications—giant or supergiant.

Once the luminosity has been calculated, we must obtain the brightness, either from a catalog of stars that lists the brightness or apparent magnitude, or by making direct observations of the star with either a photocell or photographic plate. Then, knowing both the luminosity and the brightness of the star, we can calculate its distance by means of the inverse square law given on p. 259.

Since this method of distance determination is based on a study of the spectrum, it is called the method of *spectroscopic parallax*. It has proved a very powerful tool in the astronomer's hands, although it involves certain difficulties. One difficulty is the determination of luminosity from the spectral lines. A slight error in judgment of line width and intensity means a large error in luminosity. Another difficulty lies in the fact that the inverse square law assumes that no light has been lost in transit from the star. But this is not always true, as will be seen in Chapter 14.

The determination of a star's spectroscopic parallax is limited mainly by the quality of the spectrum we can obtain. If a good spectrum can be obtained, the star is certainly bright enough for an accurate measurement to be made of its brightness, since more light is required to produce a good spectrum than to activate a photocell or to take a direct photograph.

Luminosity, Temperature, and Size

The luminosity of a star depends on both the temperature and the diameter of the star; therefore the luminosity and the temperature can be used to obtain some idea of stellar sizes for stars that are too small or too far away to be measured with the beam interferometer. If we assume that all stars radiate as nearly like a perfect radiator as does the sun, then we can determine their size by using the Stefan-Boltzmann law. We first need the luminosity of the star, which can be obtained from its brightness and distance or from its position on the H-R diagram. From this we know the total amount of energy it radiates into space. From the spectrum of the star we can learn its temperature which, with the help of the Stefan-Boltzmann law, tells us how much energy it radiates *per unit area* of its surface. If we know both how much energy the star radiates per unit area and the *total* amount of energy it radiates, then by simple division we can find the total surface area of the star. Assuming the star to be spherical (and the vast majority are spherical), we can then determine its diameter from the geometry of a sphere (Figure 10-10).

This method of determining stellar diameters is not without its difficulties. Not only is there likely to be an error in the calculation of luminosity, but also some stars may not radiate as closely to a perfect radiator as does the sun. Nevertheless, the method is an important factor in astrophysics, for it at least yields some idea of stellar sizes; and if we are to describe the stars we certainly need to know how large they are.

Mass, Luminosity, and Density

One of the most important and yet one of the most difficult characteristics to determine is the mass of a star. The only opportunity for making direct measurements of stellar mass is afforded by a star

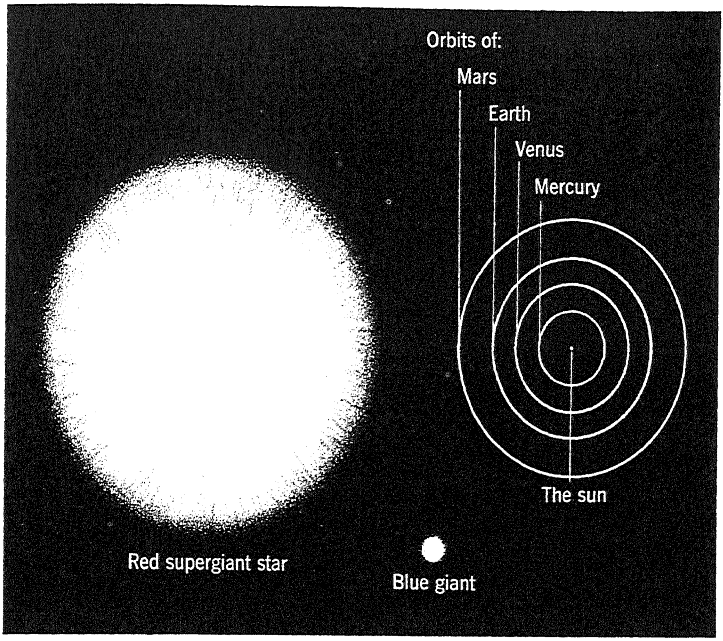


Figure 10-10 The relative sizes of some giant stars, the sun, and the orbits of the four terrestrial planets.

Jolly Green Giant

that is a member of a double-star system. Since the gravitational field about a star depends solely on its mass and since the motions of the two stars in a double-star system depend on their mutual gravitational attraction, their masses can be determined if we can describe their motions. This process is explained in Chapter 13; we may, however, discuss a closely related point at this time.

After the masses of a number of stars had been determined a close correlation between mass and luminosity was recognized. The more massive a star, in general, the more luminous it is. This *mass-luminosity relation* gives us a rough method for determining the mass of an isolated star. The mass thus determined is only an approximation, for we can see in Figure 10-11 that the stars plotted scatter about a mean curve that must be used to determine the mass. Since there is considerable deviation from this mean curve, the mass of any one star is always in doubt. When a large number of stars is used, however, the method can be employed very effectively to determine the average mass of a particular *group* of stars. For example, the masses of the main-sequence stars of each spectral type have

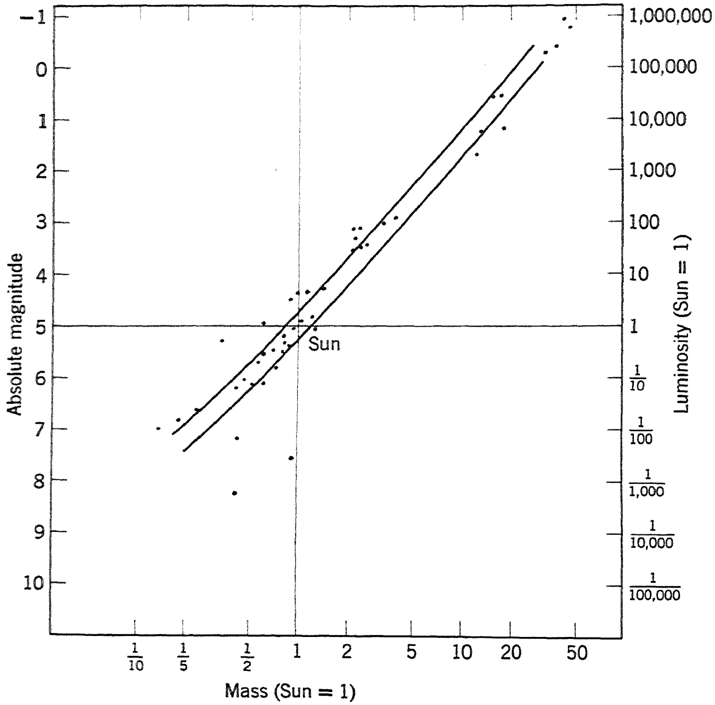


Figure 10-11 The mass-luminosity relation indicates that stars of increasing mass will be more luminous.

been determined by this method and are given in Table 10-1, from which it appears that the masses vary less than the other characteristics.

Since direct determination of mass can be made of only those stars that are members of binary systems, the mass-luminosity relation is based on such stars. If single stars differ from stars that are members

Table 10-1

Spectral type	*Mass*	*Spectral type*	*Mass*
B0	16	G0	1.0
B5	6	G8	0.9
A0	4	K0	0.8
A5	2	K5	0.6
F0	1.5	M0	0.5
F5	1.3	M5	0.2

of binary systems, this empirical mass-luminosity relation has limited value. It is known that some types of stars (for example, white dwarfs) do not follow the mass-luminosity relation and there are even special classes of binary stars that violate it. It is generally held, however, that most stars do follow this relation, which has become an important part of stellar astronomy.

The most massive star known is an O8 star which is a member of a double-star system. The name of that star is HD 47129 and each star in the system has a mass close to 50 times that of the sun. The least massive star known is less than 0.1 solar mass.

Once we know both the mass and the diameter of a star we can calculate its average density, for the density is equal to the mass divided by the volume. The densities of stars vary over a considerable range. The highest average densities are found in the white dwarfs (the subluminous stars) which are as much as 100,000 times more dense than water. A quart of material taken from such a white dwarf would weigh about 200,000 lb if placed here on the Earth. Such fantastic densities scarcely seem believable yet they have been verified both by many observations and from a theoretical point of view.

Densities as high as this can be found only in *degenerate* matter (see p. 321ff), a state of matter that can best be understood through a consideration of atomic structure. A nucleus has a diameter of the order of 10^{-13} cm, whereas an atom has a diameter of the order of 10^{-8} cm. Thus an atom's diameter is about 10^5 times that of the nucleus and its volume is 10^{15} times that of the nucleus; that is, an atom has sufficient volume to enclose 1,000,000,000,000,000 nuclei! In other words, atoms are composed mostly of empty space between very tiny atomic particles. If this space were eliminated atomic particles could be packed much more closely together, and densities could reach prodigious values.

Under the tremendous pressures encountered in white dwarfs the atomic structure does break down; the electrons and nuclei are packed together much more closely than is possible when the electrons occupy their usual energy levels. The space that occupied most of the atom has been largely eliminated and the same amount of matter that goes to make up a star like the sun, for example, can be packed into a volume less than that of the Earth.

On the other end of the scale of densities are the supergiants, which are very "inflated" stars. They have mean densities that would be considered a fairly good vacuum here on the Earth.

The question of stellar mass is very important, for on it hinges

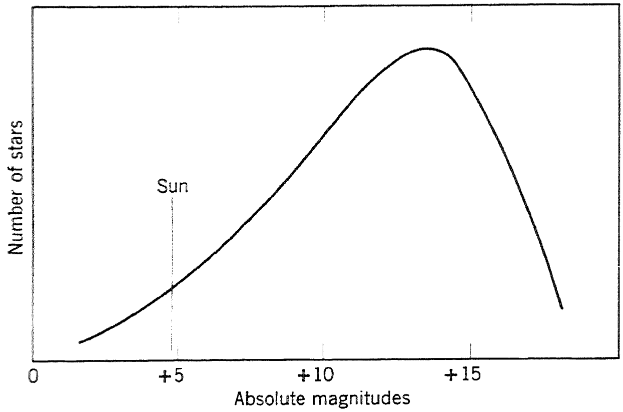


Figure 10-12 The luminosity function indicates that the less luminous stars are more numerous than the very luminous ones.

not only the process of stellar formation but also the life of the star. If stars form (as was intimated in the discussion on the origin of the solar system) from gaseous nebulae, this process of formation can be better understood if the astronomer knows the minimum amount of material necessary to form a star as well as the maximum. Furthermore, he wants to know if it is the less massive or the more massive type of star whose formation is more probable. But since the mass of an individual star can be determined directly only if it is a member of a double-star system, we must resort to luminosities to study how the stars in general vary in mass.

By making a graph of the number of stars of a given luminosity, there was formed what is called the *luminosity function*, which tells us what proportion of the stars have a given luminosity (Figure 10-12). It can be seen from the figure that the faint stars are much more numerous than the bright. Thus, the less massive stars are much more numerous.

Does this mean that they form more readily and more frequently from the nebulous material? Or does it mean that they are more numerous because they have a longer life? Houseflies are very abundant because they are born in quantities too numerous for comfort, despite the fact that they live for only a few weeks. Human beings, on the other hand, are numerous because they have a life expectancy of 65 to 70 years, even though their birth rate is so much lower than that of the fly. The answer to the question "Why are less massive stars so numerous?" may partly depend on their longer life (see Chapter 12). Whether they are more readily formed is a question that has yet to be answered.

The formulation of the luminosity function was not an easy task. Essentially it involved a stellar census; that is, the number of stars in a given volume of space had to be counted and the luminosity of each determined. We are handicapped, however, for although we may know how many of each kind of star exist in the neighborhood of the sun (within a sphere whose radius is, say, 16 light years) it is difficult to go much further. Not only do heliocentric parallaxes fail farther out but also the faint stars soon become too faint to be seen.

Consequently, the portion of the curve dealing with the low-luminosity stars was formed from the region immediately about the sun—the "local swimming hole" as the late Walter Baade, of the Mount Wilson and Palomar Observatories, so quaintly described it. Unfortunately, however, the local swimming hole contains no very luminous stars. Hence, the portion of the curve dealing with very luminous stars had to be formed from studies of a region of space beyond the reach of the method of heliocentric parallax. Their distances were obtained by the method of spectroscopic parallaxes as well as by methods to be discussed in Chapter 15. Because of these inherent difficulties, the luminosity function represents only two regions of space: within about 16 light years of the sun for the least luminous stars and beyond this up to within about 1,500 light years for the most luminous. It does not represent the entire galaxy.

The last characteristic, chemical composition, is not startling; it simply answers the question, "What are stars made of?" No matter what substances may compose it, we are fairly certain that every star is gaseous throughout, even the white dwarfs with their fantastic densities. It is possible, although difficult, to determine the chemical compositions of some of the stars' atmospheres from an analysis of their spectra. But if we want to go inside a star, we must, for obvious reasons, forego direct observation and use instead the fruits of theoretical studies.

These studies are in the forefront of astronomical thought. The conclusions reached indicate that stars begin life as balls of gas composed almost entirely of hydrogen. As they generate energy by converting hydrogen into helium, the amount of hydrogen decreases and the amount of helium increases. As they age, the helium is converted into the heavier elements. Thus a star's chemical composition depends on its age.

There are some stars, however, that have in their atmosphere an unusual abundance, as yet unexplained, of one element or another. Examples are the carbon stars of the spectral classes R and N, and the S-type stars with their relatively high percentage of zirconium and technetium. These are only some of the many deviant stars whose behaviors do not fall within the confines of this chapter's description of "normal" stars.

Basic Vocabulary for Subsequent Reading

- Absolute magnitude
- Angle of parallax
- Apparent magnitude
- Degenerate matter
- Giant star
- Light year
- Luminosity
- Main sequence
- Parsec
- Peculiar velocity
- Proper motion
- Spectroscopic parallax
- Supergiant star
- Tangential velocity
- White dwarf

For Further Reading

Kruse, W., and W. Dieckvoss, *The Stars,* The University of Michigan Press, Ann Arbor, Mich., 1957.

Page, Thornton, ed., *Stars and Galaxies,* Prentice-Hall paperback, Englewood Cliffs, N.J., 1962.

Struve, Otto, and V. Zebergs, *Astronomy in the 20th Century,* Chapters X, XI, XII, XIII, Crowell Collier and Macmillan, New York, 1962.

Barton, Randolph, "Some Effects of Stellar Proper Motions," *Sky and Telescope,* p. 4, January 1963.

Boyce, Peter B. and W. M. Sinton, "Infrared Spectroscopy With an Interferometer," *Sky and Telescope,* p. 78, February 1965.

Brown, H. R., "The Stellar Interferometer at Narrabri Observatory," *Sky and Telescope,* p. 64, August 1964.

Hack, Margherita, "The Hertzsprung-Russell Diagram Today," *Sky and Telescope,* part I, page 260, May 1966; part II, p. 332, June 1966.

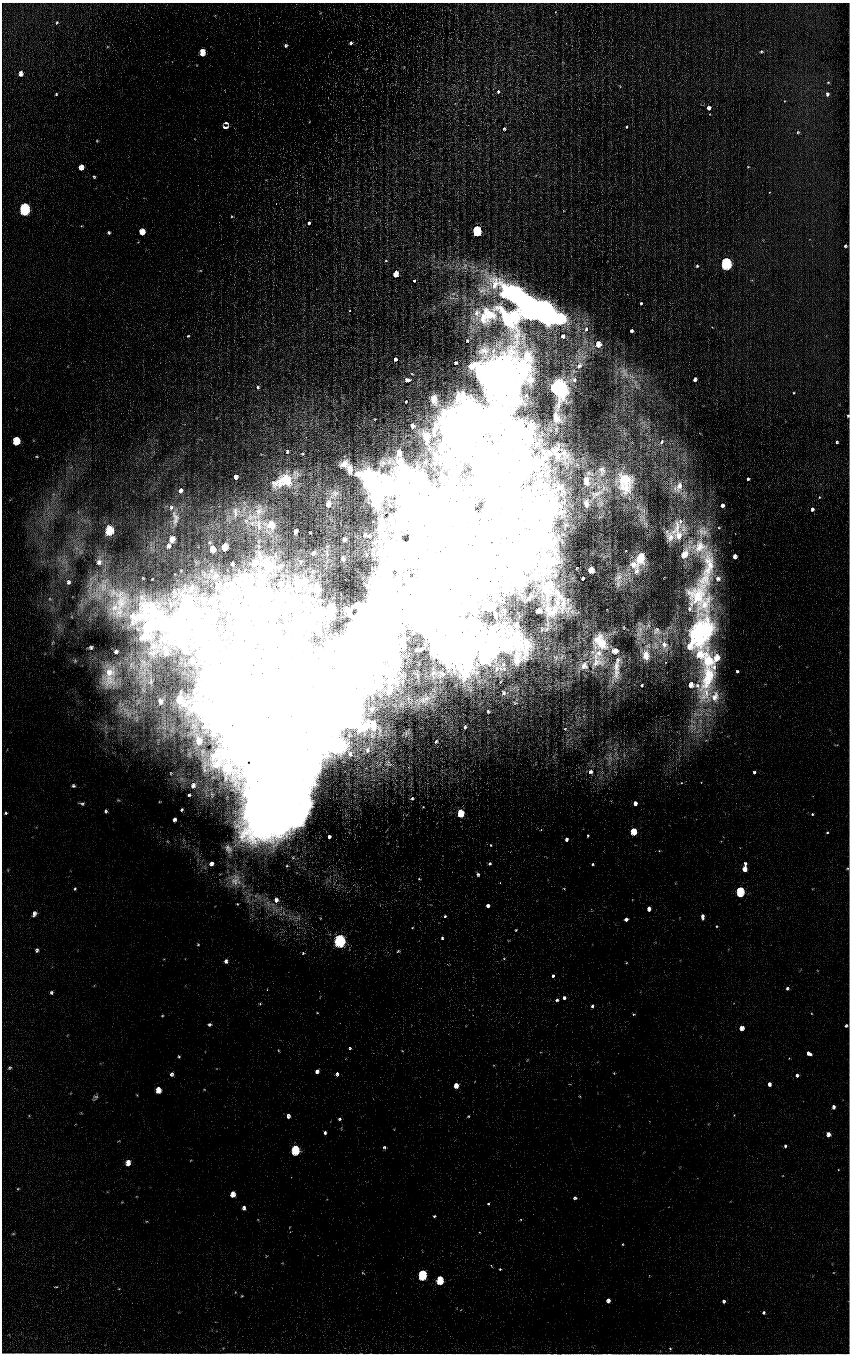
Mumford, George S., "Distance Modulus," *Sky and Telescope,* p. 274, May 1965.

Sitterly, B. W., "Symposium on Abundances of Elements in Stars," *Sky and Telescope,* p. 11, January 1965.

Wilson, O. C., "A New Scale of Stellar Distances," *Scientific American,* p. 107, January 1961.

"Extremely Cool Stars," *Sky and Telescope,* p. 195, October 1965.

"Smallest Known Star?" *Sky and Telescope,* p. 314, June 1962.



An unstable star ejected gases which now form the slowly expanding Dumbbell nebula, NGC 6853. (Lick Observatory photograph)

Chapter 11 *Nonstable Stars*

Nonstable stars, as their name implies, are stars which evidence a lack of equilibrium between gravitational pressure and gas pressure. This evidence is a change in brightness which may indicate either a periodic pulsation, and irregular eruption, or a massive and nearly self-destroying explosion. It may be that many main-sequence stars will, by the normal processes of stellar evolution, become unstable in a manner described in this chapter. For this reason, nonstable stars are of particular interest to us; and, since their study is fraught with difficulties, their problems are a challenge.

Cepheids

The most important class of these nonstable stars is called the *Cepheids*. The group derives its name from its most prominent member, Delta Cephei, a naked-eye star that misbehaves in that its lumi-

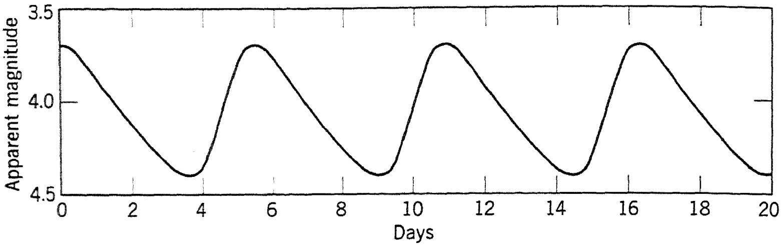


Figure 11-1 A schematic drawing of the light curve of Delta Cephei.

nosity varies in a periodic fashion. This variation of luminosity is detected by a variation in brightness or apparent magnitude. The apparent magnitude of Delta Cephei varies from about +4.3 at minimum brightness to about +3.6 at maximum. A variation of 0.7 in apparent magnitude amounts to a factor of almost 2 in the brightness. Thus, Delta Cephei at its maximum brightness is nearly twice as bright as at its minimum.

The light variation of all the Cepheids is periodic in nature, repeating itself over a certain interval of time with only small deviations from the observed pattern. When the pattern of variation for a given Cepheid has been determined from repeated observations, it is possible to predict with considerable accuracy when it will be at its maximum brightness. Indeed, it is possible to predict when it will have any given magnitude within its range of variation. Such a prediction may be made from a graph on which the variation in brightness is plotted against time. The resulting graph is called the *light curve.*

The light curve for Delta Cephei is shown in Figure 11-1. The interval of time elapsing between two successive maxima or minima is called the *period.* Delta Cephei's light curve shows that its period is close to 5½ days. A light curve made from observations over a longer time interval enables us to measure its period at 5 days, 8 hours, 46 minutes, 38 seconds. But not all Cepheids exhibit this period of light variation. The shortest period for a Cepheid is just over a day and the longest is about 50 days.

Since the periods are longer than 1 day, it is not possible to follow the variation of one entire period on any one night. The light curve, therefore, must be determined from many observations of brightness over many nights; all these observations must then be fitted onto a curve that can account for the past variations and be used to predict

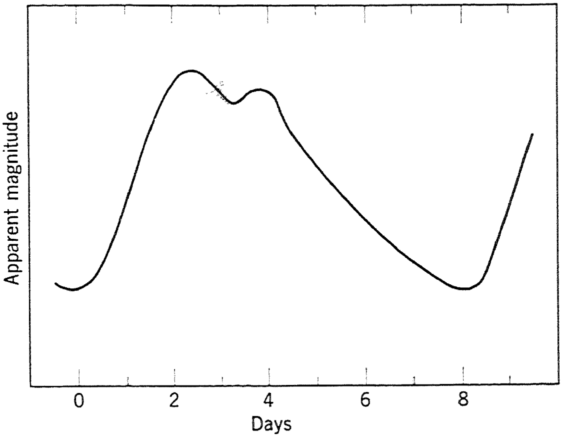


Figure 11-2 Light curves of Cepheids with periods of 7 to 8 days have a secondary maximum just after the primary maximum. These two maxima may result from two different modes of pulsation proceeding simultaneously and superimposed on one another.

future ones. Thus it may take many months to obtain the light curve of a Cepheid with a period of 45 days.

Upon examining the light curves of many Cepheids of different periods we find that these curves differ slightly in shape. Those with a shorter period have steep narrow maxima, and those with longer periods have maxima that are successively broader. Nevertheless, all Cepheids increase in brightness more rapidly than they decrease; their light curves are all asymmetrical. Those with a period of around 7 or 8 days develop a slight hump on the curve after the star has reached maximum and is on the decline (Figure 11-2). These humps repeat themselves fairly accurately each time the star goes through the corresponding phase of the period.

The cause of light variations cannot be determined from the light curves alone; for this reason the spectra of the Cepheids have been investigated quite thoroughly and with startling results. There is not only a periodic change in the spectral lines but there is also a shifting of the lines; they oscillate back and forth across a mean position. Such a shifting of lines has been interpreted as a velocity or Doppler shift. If this interpretation is correct, and all indications are that it is, then there must be some periodic motion of the star; the star must either revolve in an orbit or pulsate. Both possibilities have been investigated, and astronomers have come to the conclusion that apparently the Cepheids are nonstable stars and pulsate like balloons that are partially deflated, then inflated, then partially deflated again, and so forth. Thus, the surface facing us periodically approaches and recedes from us, thereby causing the spectral lines to shift periodically to the blue and to the red.

When we plot a Cepheid's variations in velocity against time we obtain a *velocity curve* for the star. Figure 11-3 shows not only the velocity curve but also the light curve for Delta Cephei; when both curves are shown together it becomes apparent that one is the mirror image of the other. The light curve is at a maximum when the spectral lines are shifted farthest to the blue, and at a minimum when the spectral lines are shifted farthest to the red. But the spectral lines are shifted farthest to the blue (position 1) when the star is *expanding* at its maximum rate—not when the star is largest. Correspondingly, the lines are shifted farthest to the red (position 3) when the star is *contracting* at its maximum rate—not when it is smallest. Thus the star is brightest not when it has the largest diameter, but when it is in the stage of most rapid expansion. The star is faintest not when it is smallest but when it is contracting the fastest. We conclude that its changing size is not the principal cause of the light variations.

In view of this let us investigate temperature, the other factor influencing luminosity, to see whether its behavior might not be a more important cause of light variations. From its light we can see that a Cepheid is usually hottest when it is at maximum brightness and coolest at minimum brightness. This variation in temperature is accompanied by changes in the star's color that can be determined by means of a photocell and the proper use of color filters. The star is bluest when it is hottest. As we might expect, these changes in temperature also result in changes of the star's spectral type. Maximum brightness is accompanied by an earlier spectral type. For example, a Cepheid whose spectral type is F5 at maximum is likely to have a spectral type of F9 at minimum.

Since the Cepheids are brightest at maximum temperature but at that point are of only intermediate size, we must assume that their brightness is largely dependent on their temperature. This is not really surprising if we recall that the luminosity of a star increases only with the square of the diameter (the surface area), whereas the amount of energy emitted per unit area, according to the Stefan-Boltzmann law, increases with the fourth power of the temperature (T^4). Most of the energy emitted by stars of the temperature of

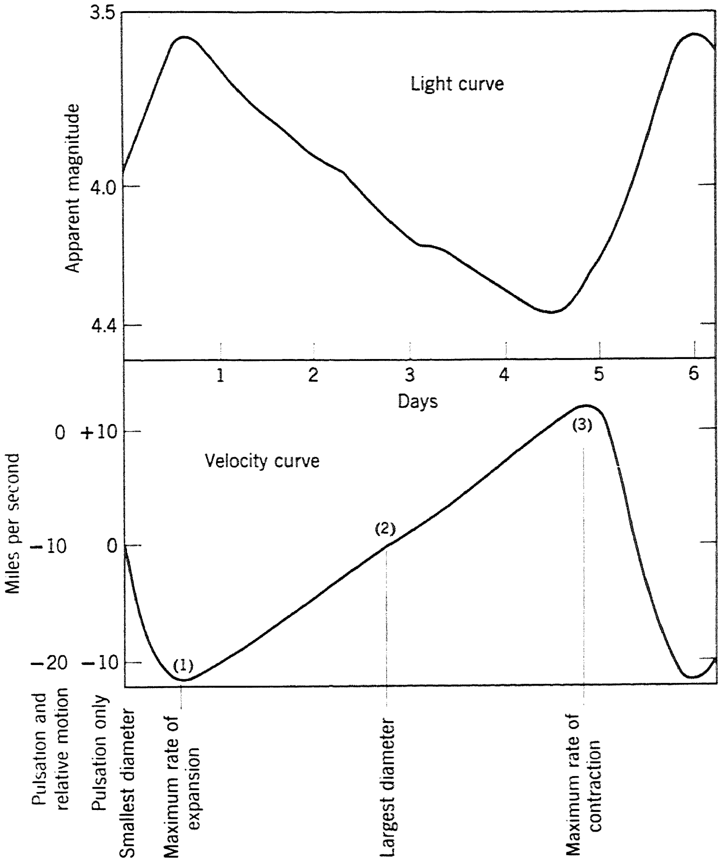


Figure 11-3 The light curve (top) and the velocity curve (bottom) of Delta Cephei. The velocity curve shows the radial velocity changes due to pulsation alone, and since the star itself is moving toward the sun with a velocity of −10 miles per second, its radial velocity relative to the sun is also given. The star is expanding whenever its radial velocity is between −10 and −20 miles per second; it is contracting whenever its radial velocity is between −10 miles per second and 0. It is expanding at its maximum rate at point (1) and contracting at its maximum rate at point (3). The star ceases expanding and begins to contract at point (2), where the Doppler shift would change from a violet to a red shift if the star were at rest relative to the sun.

Cepheids is in the visible region of the spectrum. The change in radius of a Cepheid generally amounts to about 10% of the average radius. The change in temperature is about the same, roughly 600° for a star whose temperature is 6,000°K.

The pulsation theory explains many of the observed features of the Cepheids, but it does not yet explain them all. Why the star began to pulsate in the first place is not understood completely, but we know that some unbalanced forces must have developed that caused an initial expansion or contraction. If the Cepheid began by expanding, it did so until gravitational forces took over and then it began to contract. When it had contracted to its smallest size, presumably smaller than the size from which it expanded the first time, it had compressed the gases within so much that they forced it to expand again and repeat the cycle. The star continued to pulsate at a rate that depended on its size and mass.

Period-Luminosity Curve

Even though the cause of pulsation remains unknown, the Cepheids, because of their light variation, constitute one of the classes of stars most important to the astronomer. Since astronomical distances are so great, and methods of measuring them are so difficult, astronomers are always looking for new and better ways to verify and extend our present knowledge of these distances. The Cepheids offer a unique method, for their luminosity and period are related so that once the period of a Cepheid is determined, its luminosity can easily be estimated. This relation, called the *period-luminosity relation,* was originally found by studying the Cepheids in the *Small Magellanic Cloud,* a large aggregation of stars visible as a faint patch of light in the southern hemisphere not too far from the South Celestial Pole.

In 1912, Henrietta S. Leavitt, of the Harvard College Observatory, plotted the periods of the Cepheids seen in the Small Magellanic Cloud against their brightness and found the relation shown in Figure 11-4: the longer the period, the brighter the star. It was realized that the cloud has a very small diameter compared to its distance from the sun and consequently all the stars in the cloud could be considered as being the same distance from the sun without introducing much error into measurements that compared the luminosities of these stars. Thus the period-*brightness* curve discovered by Miss Leavitt should mean that a period-*luminosity* curve exists, and if the luminosity of even one Cepheid could be determined along with its period, the luminosity of each Cepheid in the Small Magellanic Cloud could be determined by reference to this standard Cepheid. That is, the scale on the left of Figure 11-4 could be converted from apparent magnitude (brightness) to absolute magnitude (luminosity). If the luminosity of each Cepheid were known, the distance of the cloud from the sun could be determined by the inverse square

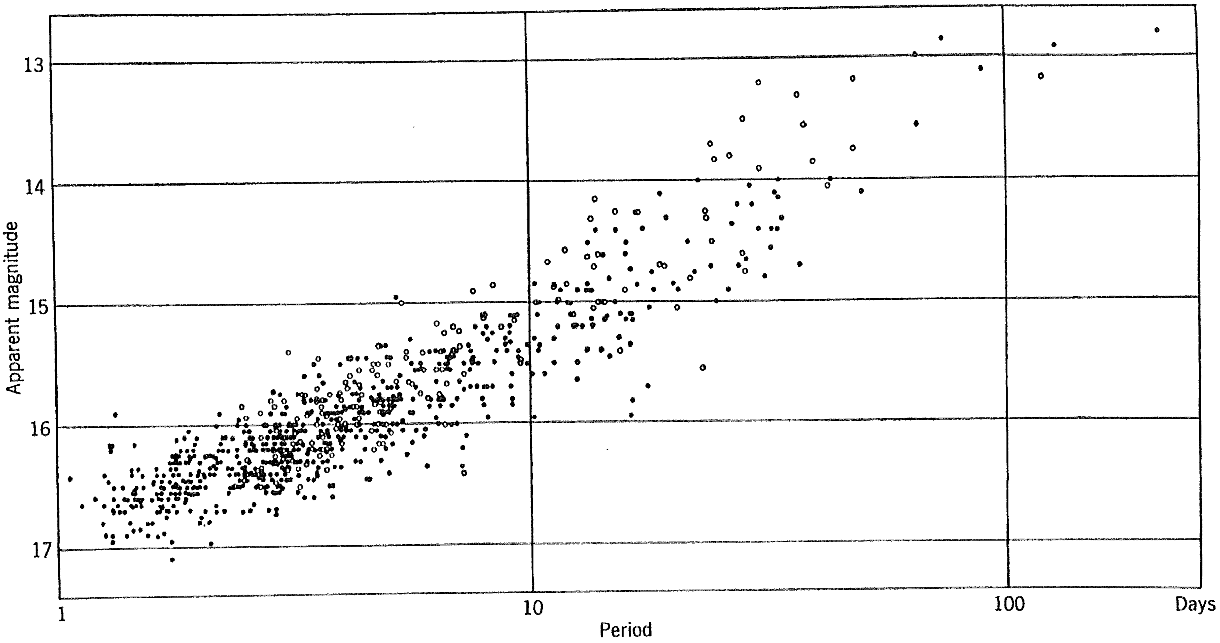


Figure 11-4 The period-luminosity relation for the Cepheids of the Small and the Large Magellanic Clouds. (Modified by permission from *Astrophysics*, edited by Hynek, Copyright 1951. McGraw-Hill Book Co.)

law, as could the distance of *any* Cepheid in the sky whose period was known.

Unfortunately, there is no Cepheid close enough to the sun for a heliocentric parallax determination. The problem of finding the luminosity of a Cepheid or group of Cepheids took on major interest for astronomers. The explanation of the method used to find their absolute magnitudes will have to wait, however, until Chapter 15. But needless to say, once the absolute magnitudes of a number of Cepheids were finally determined, the distance of every Cepheid seen in the sky could be calculated. It was mainly through this discovery that the astronomer began to realize that the universe is really much larger than he had previously thought, and he began to speak in terms of millions of light years. Before that time (in the mid-1920's) it was not known with any certainty whether the so-called "spiral nebulae" were a part of our galaxy or were outside and enormously farther away and larger.

As fruitful as this discovery has been, all the distances based on the Cepheids had to be revised. In the early 1950's, Baade realized that stars can be divided into two basic types which he called Population I and Population II. Some of the Cepheids are Population I and others (called W Virginis stars) are Population II, and each has a different period-luminosity relation. His discovery was based on observations of the Andromeda galaxy, a neighbor of our own galaxy.

It developed that Population I Cepheids are about 1.5 magnitudes more luminous than Population II Cepheids and the Cepheids studied by Miss Leavitt are actually more luminous than was earlier assumed. If they are more luminous, they must be farther away than the original calculations had led astronomers to believe: a magnitude difference of 1.5 amounts to a factor of almost 2 in distance. In fact this discovery indicated that the universe outside our own galaxy is twice as big as we had previously thought.

A neat clean period-luminosity relationship seems almost too good to be true. And so perhaps it is. Recent studies indicate that although there is a relationship between the period and the luminosity of Cepheids, it is complicated by the surface temperature of the stars. Thus, two stars with the same period but different surface temperatures will have slightly different luminosities. The result is that the period-luminosity relationship is spread out into a band about 1 magnitude thick, for both Population I and Population II Cepheids.*

---

* The nature of stellar populations will be discussed in more detail later.

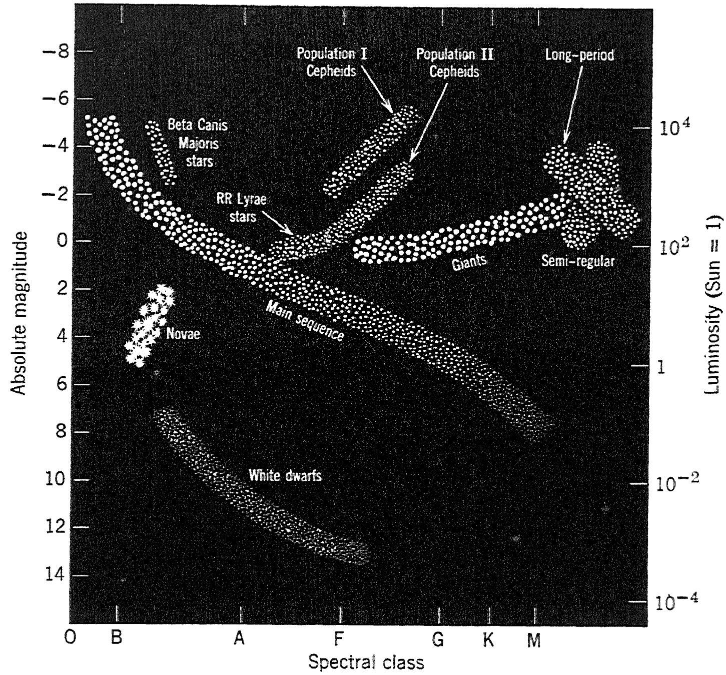


Figure 11-5 A schematic H-R diagram showing the location of many of the nonstable stars in relation to the main sequence and the giants.

Now that the luminosities of the Cepheids are known, it is instructive to plot them on the H-R diagram. (Since both their luminosities and their spectral types vary, the average of the extremes is used for such purposes.) Figure 11-5 shows the position of the Cepheids (along with some other variables to be discussed in this and later chapters) on the H-R diagram. The Cepheids' luminosity places them in the giant and supergiant class, which is fortunate because they are visible even at great distances from the sun.

The H-R diagram shows also that among the Cepheids there is a relationship between spectral type and luminosity; the later the spectral type, the more luminous the star. An F-type Cepheid has a luminosity of about 100, and a G-type has a luminosity of about 10,000. At first glance this may seem contrary to the Stefan-Boltzmann law until we recall that this law only refers to luminosity *per unit area.* Thus, the later-type Cepheids must be very much larger than those of early type.

There is, then, a very definite relationship between the period of a Cepheid and a number of its other characteristics. An increase in period is accompanied by: (1) an increase in luminosity; (2) an increase in diameter; (3) a decrease in temperature; (4) a larger range in temperature variation; and (5) a larger range in luminosity variation. That all of these are so related is significant and must depend on the masses and sizes of the various Cepheids.

RR Lyrae Stars

There are other stars related to the Cepheids in that they, too, appear to pulsate, and yet they differ enough to be classified separately. The first of these to be discussed are called the *RR Lyrae* stars after one of their members. Because they are often found in globular clusters (see p. 360ff) they are sometimes called cluster-type variables. These stars belong to Population II and have periods ranging from about 1½ hours to a little over 24 hours. The longer-period RR Lyrae stars blend into the type II Cepheids, since there are several stars that could fall into either group. The most luminous RR Lyrae stars have an absolute magnitude of 0.0 (close to 100 times the luminosity of the sun), and the least luminous have an absolute magnitude of +1.2 (Figure 11-5); the average is about +0.6. As with the Cepheids, the range in luminosity for a given period is attributed to differences in surface temperature.

The greater luminosity of the Cepheids results in their being seen at much greater distances than the RR Lyrae stars. The Cepheids can be used as distance indicators outside of our Milky Way Galaxy; RR Lyrae stars can be so used only inside our galaxy and its immediate environs. But, despite problems yet to be resolved, both groups of stars have been indispensable links in the measurement of distances within our galaxy and throughout the visible universe.

As do the Cepheids, most of the RR Lyrae stars increase in brightness rapidly and decrease less rapidly, but the symmetry of their light curves varies over the range of their periods. The short-period RR Lyrae stars are likely to have more symmetrical light curves, increasing and decreasing in brightness in about the same time. Light variation is also indicative of a temperature change. Most RR Lyrae stars are early A-type stars at maximum brightness, and become early F-types at minimum brightness.

Spectral classification of these stars, however, is rather difficult because of certain anomalies in their spectra. One such anomaly is that at certain phases during the period of an RR Lyrae star, double hydrogen spectral lines are observed. That is, each line that normally appears as one becomes two. A doubling of spectral lines can be explained by assuming that there are two layers of gas expanding and contracting out of phase with each other, each causing a different Doppler shift.

This speculation has been further strengthened by the observation of hydrogen emission lines at a time when the two layers of gas should collide. That hydrogen gas should emit bright lines at this time has been verified by aerodynamical experts who have recorded similar emissions from gases in powerful shock tubes.

RR Lyrae stars were one of the stepping stones toward the identification of Population I and II stars and, consequently, to the new distance scale. On the old distance scale the Andromeda galaxy was considered to be about 750,000 light years from the sun. Assuming that the Andromeda galaxy is similar to our own, if it were 750,000 light years away, Baade knew that he should be able to detect RR Lyrae stars in its midst; he became suspicious, however, when not a single RR Lyrae star could be detected. Detailed photographs and studies revealed that the nucleus of Andromeda contained no blue giant stars, only red giants; in its spiral arms Baade found blue giants and relatively few red giants. Their preferred location indicated to him that these stars were different in some fundamental respect, a difference he indicated by naming the blue giants Population I stars and the red giants Population II. It is this work that first pointed the way to a division of stars into various Populations. The knowledge that the Cepheids are among the stars thus divided enabled astronomers to redetermine the distance of the Andromeda galaxy at about 2.2 million light years, too far for its relatively faint RR Lyrae stars to reveal themselves on photographs.

Long-Period Variables

Long-period variables form another group of nonstable stars. The stars in this group are red giants and have periods ranging from about 90 to over 700 days, with the majority around 300 days. Their light curves do not repeat themselves as accurately as those of the Cepheids or the RR Lyrae stars, and there are irregularities not only in the shape of such a light curve but also in the period of the star. The reasons for these irregularities are not known, but it has been suggested that they might be connected with the large size and low density of these red giants, whose diameters are several

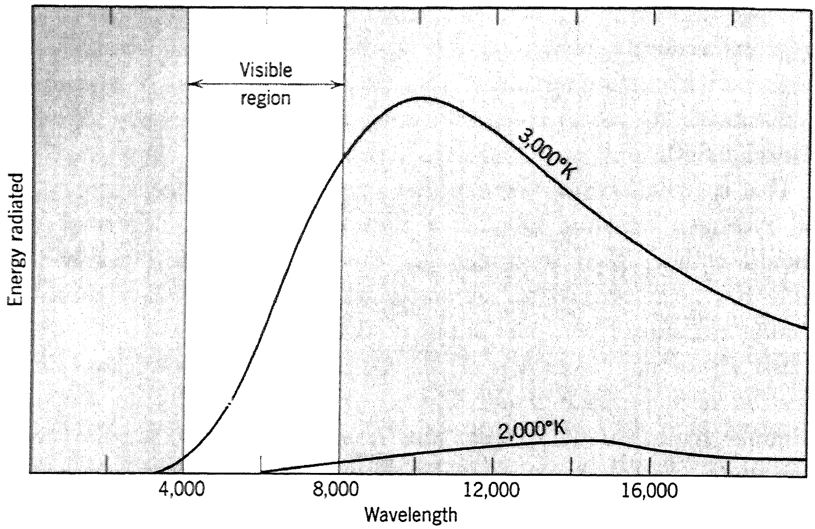


Figure 11-6 The radiation curves of stars at 2,000° and 3000°K showing the large difference in the amount of visible light each radiates.

hundred times that of the sun and whose average densities are only about 1/1,000,000 the density of water.

The range of luminosity for the long-period variables is about 5 magnitudes, considerably larger than for the Cepheids. Thus a star at maximum will be about 100 times brighter, visually, than at minimum. We know of at least one star of this type whose luminosity varies by a factor of 4,000.

This large variation in luminosity is partly accounted for by the fact that the long-period variables are relatively cool stars, many of them M-type stars having temperatures of around 2,000° to 3,000°K. According to Wien's law, a star as cool as this will emit most of its radiation in the infrared (maximum intensity between 10,000 and 15,000 Å). Hence any variation in temperature will affect the visible region of long-period variables much more than would normally be expected in stars having a temperature variation of only around 400°K.

The effects of temperature variation can be seen clearly in Figure 11-6, where the radiation curves are drawn for temperatures of 2,000° and 3,000°K. The area under the curve represents the amount of energy radiated. The portion of this area falling within the visible region is much smaller for a star with a temperature of 2,000°K than for one with a temperature of 3,000°K. This is why a small change in temperature results in a large change in the amount of visible light. If the total energy—visible and invisible—emitted by these stars is observed, the variation is only about one magnitude.

The long-period variables differ from the Cepheids in other ways. When the light curve of a long-period variable is compared with its velocity curve, we find that the time of maximum brightness often coincides with the maximum rate of *contraction* and not, as for Cepheids, with the maximum rate of expansion. However, the velocity curves of long-period variables are difficult to determine, for these stars exhibit only small changes in velocity and these changes are not consistent within the whole group.

Since the long-period variables are so cool, molecular bands appear in their spectra; in fact, a number of these stars are classified as R-, N-, or S-type stars. During minimum temperature the intensity of their molecular bands increases to such an extent that they effectively block out a considerable part of the visible spectrum. These stars therefore emit even less visible light during minimum temperature; this combined with the effect resulting from the shifting of the radiation curve causes even more extreme variations in luminosity. The long-period variables have been likened to a star surrounded by a molecular cloud or cloak which obscures the star beneath. Their spectra will often show bright lines near the time of maximum luminosity, but these lines fade as the luminosity decreases. It has been suggested that these bright lines may be caused by a series of shock wave phenomena that move out from the center of a star. The energy thus carried causes the bright hydrogen lines.

Several long-period variables are visible to the naked eye. The most notable is Mira in the constellation of Cetus, which on occasion has been known to be a second-magnitude star at maximum and is then easily visible. But Mira drops out of sight eventually to become a ninth-magnitude star at minimum, returns to visibility and to its maximum during its period of 331 days. It is visible to the naked eye for a total of about 6 months of this period.

The irregularities in the variations of long-period variables are to a certain extent characteristic of all the red giants. In fact, red giants with large irregularities are often called *semiregular* variables. Their periods (or cycles, since "period" implies regularity) seem to average around 100 days, with another group averaging around 350 days. Thus, there appear to be two types of stars composing the semiregular

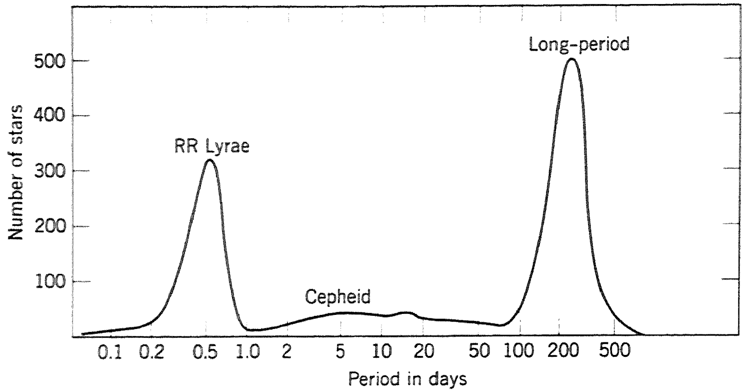


Figure 11-7 **The distribution of pulsating variables.**

variables. Their spectra classify them all as M-, N-, R-, or S-type stars, and most of them show bright lines in their spectra. They cannot be comfortably classified in either Population I or II; it has been suggested that they may form a bridge between the two populations.

The distribution curve of stars that pulsate in a more-or-less regular fashion is shown in Figure 11-7. The long-period variables are the most common, with the RR Lyrae stars next. The actual ranges of periods for the groups overlap, and, as we have seen, an arrangement according to period alone does not give us an arrangement according to other properties of these stars. It does yield the relative number of stars of each period, however, and this will be significant for our discussion of a star's life history (Chapter 12).

Novae

Some obviously nonstable stars erupt or burst with a display of energy so large it is difficult to imagine. Such outbursts have been noted in history when a star so increases in brightness that it becomes easily visible where no star was seen before; it is called a *nova* (new), even though we now know that novae are not new at all.

Many novae have been recorded in the history of Western civilization. One of the more notable, which occurred in 134 B.C., stimulated Hipparchus to compile his catalog of the stars by position and apparent magnitude. Another, in November 1572, was observed by Tycho Brahe, and is called Tycho's nova. This nova at its brightest appeared as bright as Venus and could therefore be seen during the daytime. Notable, too, was Kepler's nova (1604) which looked as bright as Jupiter.

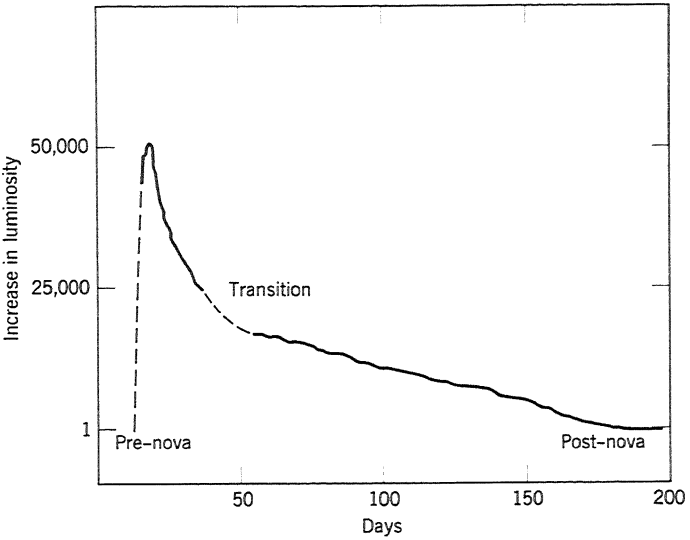


Figure 11-8 A schematic light curve for a typical nova.

Bright novae which have occurred since the development of modern equipment have been observed as thoroughly as possible. Nova Aquilae (1918) was nearly the brightest star in the sky and is the first nova which had been recorded before it became a nova. Direct photographs show that Nova Aquilae was perhaps an A-type star of variable brightness before it became a "new" star. Thus, novae are really existing stars that increase in brightness many thousands of times. After such an outburst of energy the brightness of a nova will gradually decrease until it returns to its former relatively insignificant state.

Figure 11-8 shows the light curve of a typical nova. The intensity of light may increase up to 160,000 (although the average is closer to 50,000) times its original brightness in a matter of hours. The exact time is difficult to determine, for a nova is not noticed until it has already achieved a brightness sufficient to distinguish it. After achieving this peak it begins to wane, at first rapidly and then, after a transition phase, more slowly, taking months or even years to return to its original brightness.

The transition phase between the initial rapid decline and the ensuing more gradual decline is very often observed as fluctuations of light, as if the star were pulsating. But a single and greater decrease in brightness has been observed during this phase in other novae, which then brighten again before declining gradually to their original magnitude.

That such an outburst of energy would be accompanied by changes in the spectrum of a nova seems a foregone conclusion, and indeed, all novae seem to go through nearly the same spectral changes. During the rapid increase in brightness the spectrum is usually that of a late B- or early A-type star, although it contains emission lines, something that normal stars of these spectral types do not manifest. The absorption lines are considerably broadened and displaced toward the violet, indicating that the region of the star facing us is approaching and turbulent. At maximum brightness, the star apparently cools, for the spectrum becomes that of a slightly later-type star, and the emission lines begin to fade.

As the light begins to fade, the spectrum becomes similar to that of a late F-type star; it develops very broad emission bands, each with an absorption line on its violet side. Eventually the absorption lines fade and the spectrum becomes a continuous spectrum with broad emission bands superimposed on it. Then, long after the star has returned to its normal brightness, the emission lines begin to fade and the spectrum may consist only of a continuous spectrum with a very few faint absorption lines.

The idea of a stellar explosion best explains these changes in the spectrum, as well as the initial phenomenal increase in luminosity. Since a star continues to shine many years after the explosion with nearly the same luminosity as before, the explosion is apparently a surface phenomenon. The star literally "blows its top." Let us now interpret the spectral changes in the light of such an explanation.

As the star's surface begins to expand, the spectral lines are shifted far to the violet. For Nova Aquilae this shift amounted to a velocity of 1,100 miles per second. The widening of the lines may result from the turbulence to be expected in such an explosion. As the shell of expanding gas grows larger, it begins to produce emission lines since it is a gas under very low pressure hot enough to emit its own light. Only that portion of the gaseous shell immediately between the star and the Earth (region *a* in Figure 11-9*a*) continues to produce absorption lines in the star's spectrum observed from the Earth. The rest of the shell (regions *bcd* and *b′c′d′*) produces bright lines in spectra recorded on the Earth.

That portion of the expanding shell immediately between the star and the Earth not only causes absorption lines but is also traveling

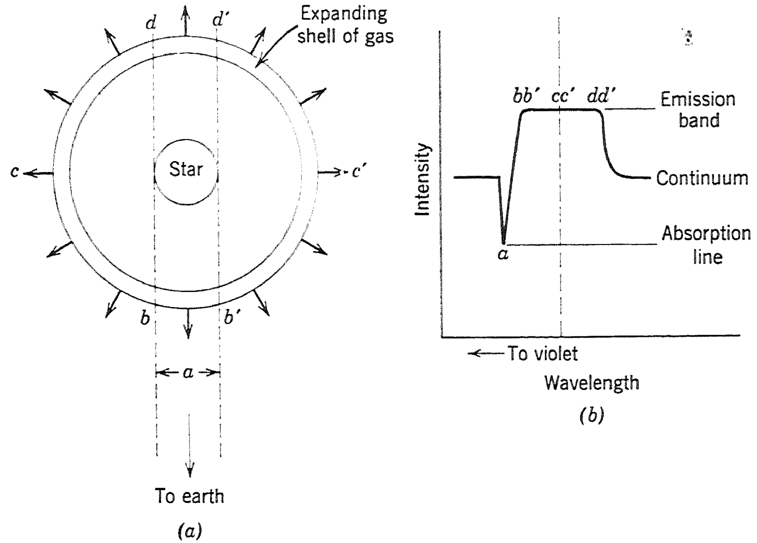


Figure 11-9 The expanding shell of gas about a nova (*a*) gives rise to a bright band in the spectra (*b*) with an absorption line on its violet edge.

toward the Earth faster than any other portion. This results in having those absorption lines shifted farther to the violet than the lines originating in any other region. The regions marked *b* and *b′* emit bright lines; they do not cause any absorption lines because they do not lie between the star and the Earth. These regions are also traveling toward the Earth, but more slowly than region *a;* consequently the spectral lines they emit are not shifted as far to the violet as the absorption lines from region *a*. Regions *c* and *c′* are traveling at right angles to the line of sight; as a result the bright lines they emit are shifted neither to the violet nor to the red unless the whole star is moving toward or away from the Earth. Since the shell is very transparent to its own light, we receive light emitted by the regions marked *d* and *d′*. But these regions are traveling away from the Earth and their bright lines are therefore shifted to the red.

The net result of these factors is an emission band (Figure 11-9*b*) given off by the shell from *bcd* and *b′c′d′*, and all the intervening regions. If these regions are considered in succession, the spectral lines they emit are shifted ever less toward the violet as we approach *c* and *c′*, from which point on they become shifted even more to the red. The blending of all these immediately adjacent bright lines produces the solid band. Again, it is region *a* alone that causes the absorption line that appears on the violet side of each emission band.

As the shell progressively expands, the spectrum changes noticeably, going through a stage where it includes some of the same lines as the aurora spectrum of the Earth's upper atmosphere. Later the nova spectrum contains bright lines that are seen in the spectra of gaseous nebulae.

During this explosion the photosphere is not quiescent. Indeed, the suggestion has been that it also expands until the star reaches maximum luminosity, at which time it starts to fall back into the star. The gas above the photosphere, having achieved escape velocity, would continue to expand and form the shell. That novae do eject gases is confirmed by photographs of Nova Persei (1901); one of these, recently taken at the Palomar Observatory, is shown in Figure 11-10.

When the shell of expanding gas is visible it gives us a method for finding the distance of the nova. From the Doppler shift of the absorption lines we can determine the velocity with which the shell is expanding. Then, when after a number of years the shell becomes visible, we can compute its actual radius in miles because the radius must equal the velocity of the ejected material multiplied by the interval of time between the explosion and the appearance of the observed shell. Since we now know the linear diameter of the shell we can calculate its distance by simply measuring its angular diameter.

By this method astronomers have been able to determine the absolute luminosity of some of the novae. Nova Aquilae (1918) had a luminosity of about 40,000 times that of the sun. Before it became a nova it had been only slightly more luminous than the sun. Such an increase in luminosity obviously required the sudden release of a tremendous amount of energy.

Supernovae

Novae, on the average, achieve an absolute magnitude of about —6 or —7 at maximum. This is, momentarily at least, very bright indeed, but there are other stars which in exploding achieve an absolute magnitude as great as —13 or even —16. In other words, an ordinary nova may achieve a luminosity of about 50,000 suns but the star that reaches an absolute magnitude of —16 achieves a luminosity of about 200 million. Stars that reach this latter, far greater magnitude are called *supernovae*. A supernova indeed, when a star

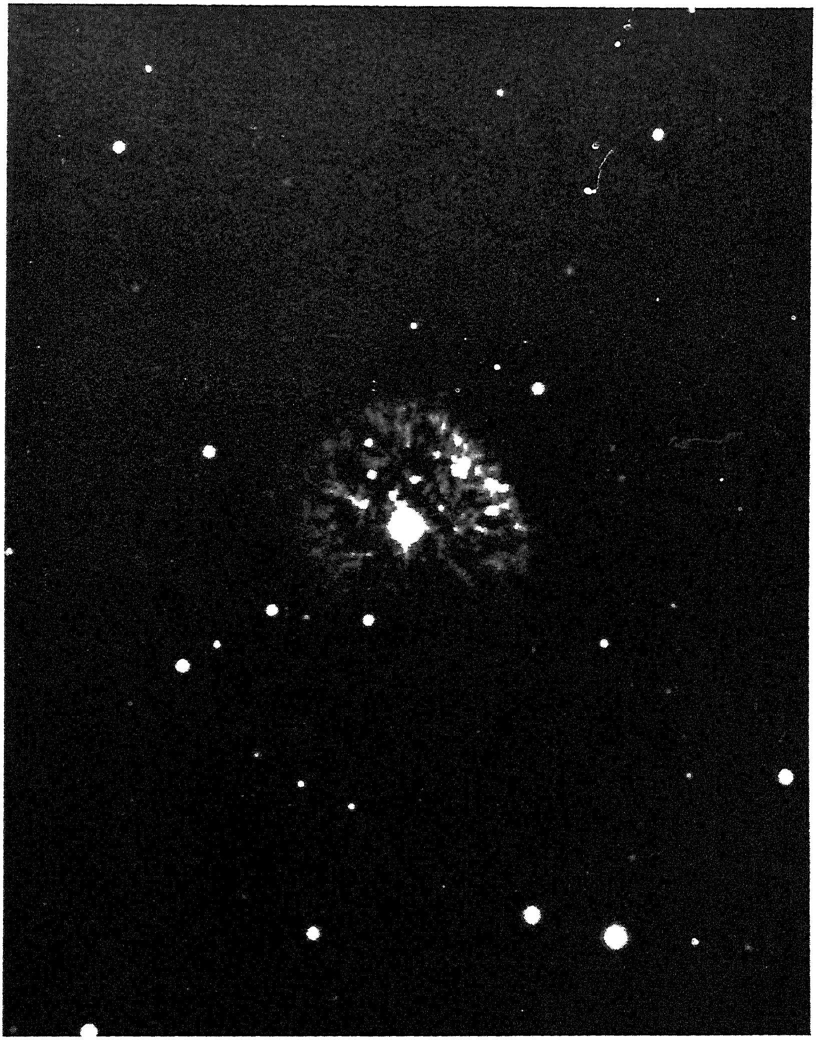


Figure 11-10 The expanding nebulosity about Nova Persei (1901). (Mount Wilson and Palomar Observatories)

gives off as much light at maximum as 200 million suns! This amount of light may be similar to the amount given off by the entire galaxy in which the star resides, and represents a positively staggering expenditure of energy. Certainly an explosion of this sort must leave a star somewhat the worse for wear.

The Crab nebula (Figure 11-11) is the result of such an explosion

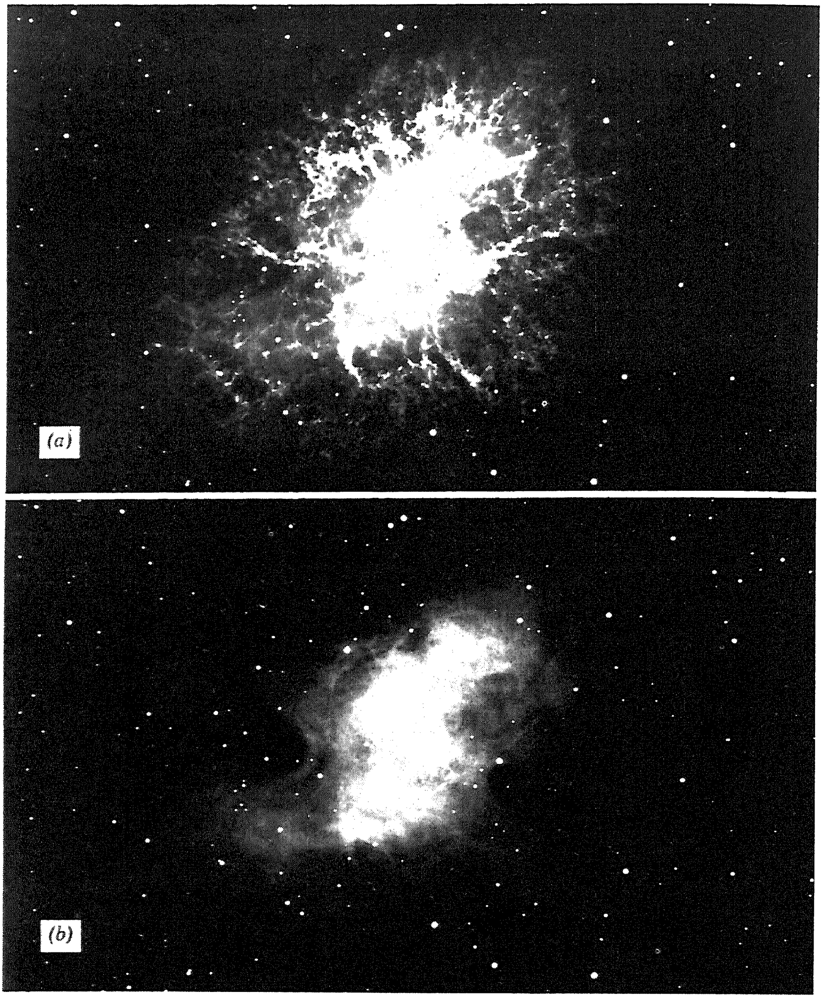


Figure 11-11 The crab nebula as seen (*a*) in the red light of hydrogen and (*b*) in green light. (Mount Wilson and Palomar Observatories)

observed in 1054. The connection between the Crab nebula of today and the supernova of more than 900 years ago is quite definite. We can observe and measure the rate of expansion of such a nebula by two methods. The first is to determine the radial velocity from the Doppler shift of a portion of the nebula that is approaching us. This gives us a linear velocity of expansion of 800 miles per second.

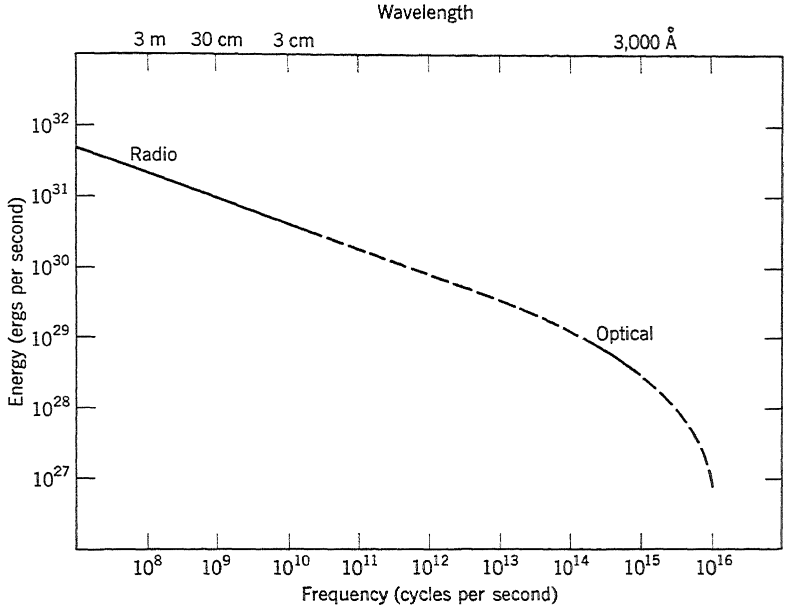


Figure 11-12 The continuous spectrum of the Crab nebula extending from the radio to the optical region. Solid lines represent observed regions. (After Steinberg, J. L., and J. Leqeux, *Radio Astronomy*, McGraw-Hill Book Co., New York)

We can also measure the angular rate of expansion by taking two pictures separated by a considerable interval of time. The combination of these two observations tells us that the nebular gas is about 4,000 light years away and is expanding at a rate such that it should have begun expanding about 900 years ago. The exact date was derived from the chronicles of ancient Chinese astronomers who recorded a "guest" star of the brightness of Venus in the celestial region now occupied by the Crab nebula. This guest star, which was also recorded by the Japanese, was visible during the daytime for 23 days and at night for more than a year.

Pictures of the Crab nebula, taken in different colors with color filters and photographic plates sensitive to the colors transmitted by each filter, reveal that the gaseous filaments are rather complex. The photograph taken in red light is centered on the $H\alpha$ line of hydrogen and it appears that the nebula's complex structure is caused principally by gases emitting this spectral line. The photograph taken in green light not only reveals that there is an amorphous structure in the center but also that the light emitted by this amorphous structure is polarized (see p. 15f).

The nebula also emits strongly in the radio region of the electromagnetic spectrum. But its radio emission is not *thermal;* that is, the intensity of radiation at each wavelength does not correspond to Planck's law of radiation (see p. 206). The amount of energy emitted decreases with decreasing wavelength, so that more radio energy is emitted than microwave energy, and more microwave energy is emitted than infrared, etc. (Figure 11-12). This observation, coupled with the fact that both the light radiation from the amorphous gas and the radio-energy radiation are polarized, leads to the conclusion that this is what we call *synchrotron radiation.*

In a synchrotron (a high-energy particle accelerator) charged particles are propelled at speeds very close to that of light, and these particles are forced by magnetic fields to follow a curved path. Moving at such high speeds in a magnetic field results in the emission of energy. The greater the speed and the stronger the magnetic field, the more energy these particles will emit; but as the speed and magnetic field are increased, the wavelength of the radiation decreases. Synchrotron radiation is polarized and corresponds to the radiation we receive from the Crab nebula. We conclude, therefore, that in the Crab nebula electrons (electrons emit synchrotron radiation easier than protons) are traveling at speeds close to that of light in a magnetic field. These electrons must move in corkscrew paths about a given magnetic line of force.

The star that remains in the center of all this gas appears to be a white dwarf, although this is not certain. The fact that something remains at all indicates that even the explosion of a supernova does not destroy the star. It does seem fairly certain, however, that much more gas is thrown off from a supernova than from a nova.

Bolstered by the successes of theories and observations of the Crab nebula, astronomers have searched the records and the skies for other supernova remnants. Their efforts have been rewarded. They have found the remnants of Tycho's nova, of Kepler's nova, and many others. The most intense discrete radio source in the sky, Cassiopeia A is now known to be a remnant of a supernova explosion which occurred about the year 1700 (Figure 11-13). There is no reference in historical records to a nova at that time, and the records of those years are fairly adequate, but it is a distant object and behind a good deal of interstellar material, so it may not have been obvious

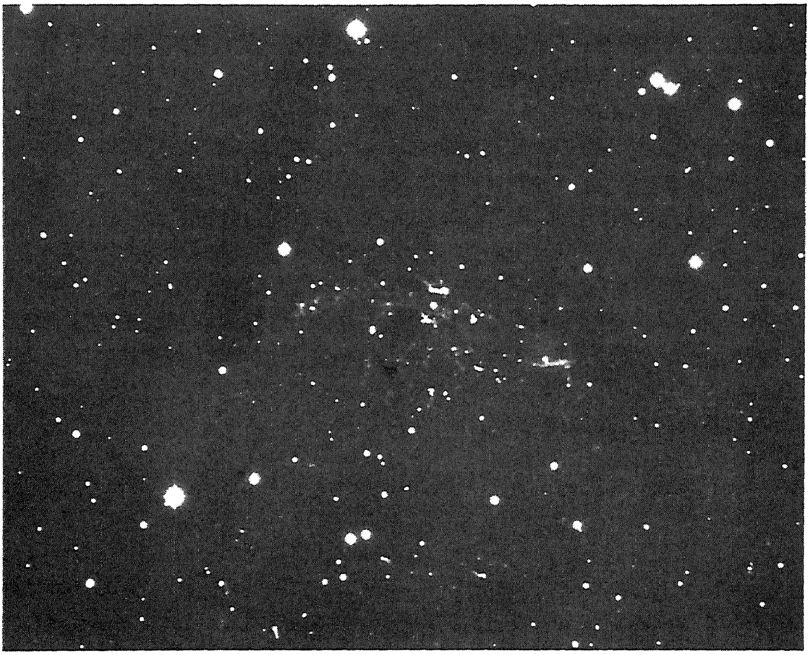


Figure 11-13 Cassiopeia A, a supernova remnant, is the strongest discrete radio source in the sky, but it is scarcely visible in this photograph made with the 200-in. telescope in red light. (Mount Wilson and Palomar Observatories)

in the nighttime sky. At the present, fragments of the nebula are expanding at the staggering velocity of 4,600 miles per second. Just why this velocity of expansion is so much greater than that of the Crab nebula is not known. At a distance of 1,000 light years it is estimated that an amount of material equal to the sun's mass was ejected during the explosion.

The Loop (Veil) nebula in Cygnus is another nonthermal source of radio energy, and it, too, is a remnant of a supernova (Figure 11-14). Material in the Loop nebula is expanding at the rate of only 56 miles per second. Presumably, its expansion has been slowed by the interaction of the gases with interstellar gases. The center of this nebula is 240 light years away, and its diameter is 12 light years. Therefore, this explosion must have taken place thousands of years ago.

It is not yet clear just why the electrons in these supernova remnants have such high velocities. It may result from the original explosion, or it is possible that electrons are being accelerated by mechanisms

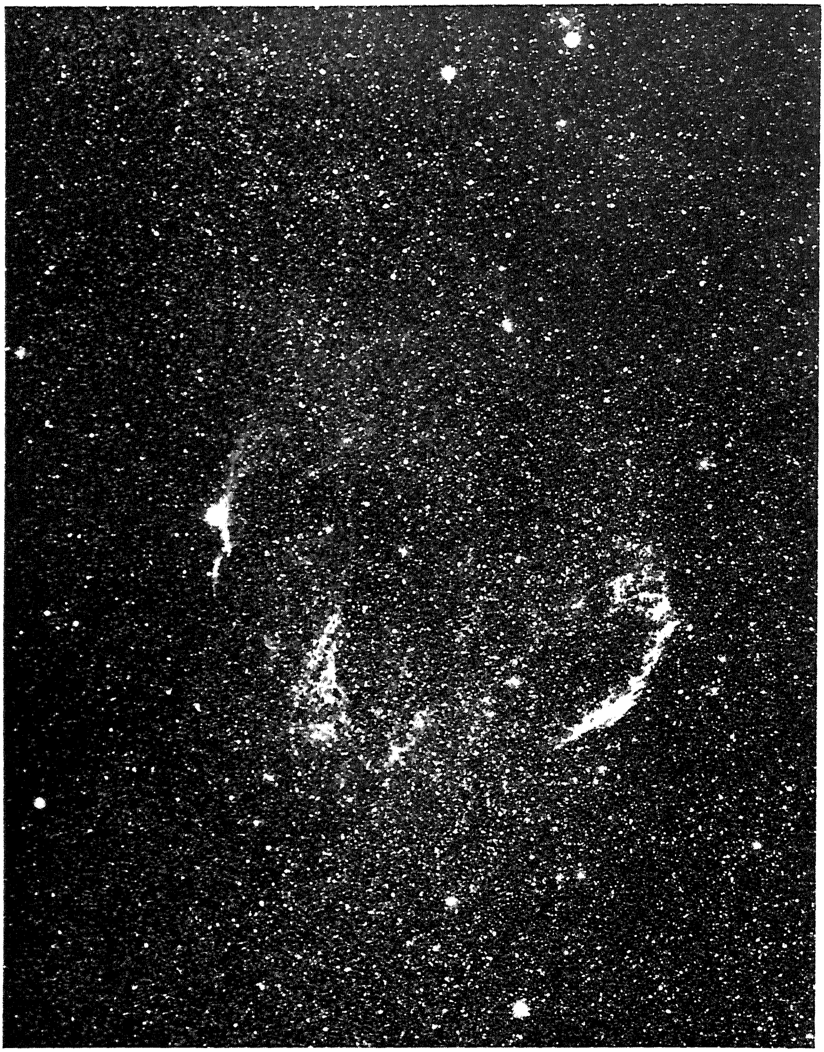


Figure 11-14 The Loop nebula in the constellation Cygnus is the result of a nova that exploded many thousands of years ago. (Mount Wilson and Palomar Observatories)

not yet understood. In any event, every nonthermal source of radio energy in our galaxy is immediately suspected of being a supernova remnant.

The Crab nebula also emits X-rays. Under the direction of Dr. Herbert Friedman, the U.S. Naval Research Laboratories in White Sands Proving Grounds, N.M., have sent rockets above the atmosphere

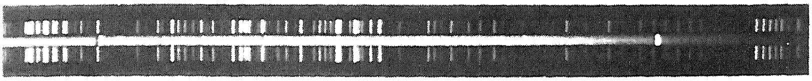


Figure 11-15 The spectrum of P Cygni showing the emission bands with an absorption line on the violet side of each. (Yerkes Observatory)

with X-ray cameras and have detected discrete sources of X-ray in the sky. The strongest of these sources lie close to the plane of the Milky Way, so these, at least, are presumably within our galaxy, although other sources are likely outside our galaxy. One difficulty is that X-rays cannot be focused, so details about the X-ray sources are hard to get. The Crab nebula, however, lies within the region of the sky through which the moon passes and so it is possible to obtain some idea of the extent of the X-ray source by observing the Crab nebula during the period of time that the moon's disk moves in front of and covers the nebula. Once occulted, no X-rays will reach us from the moon. Even though the timing and observation are difficult to make, this observation has been made by sending a rocket above the Earth's atmosphere during just the few minutes of onset of a lunar occultation of the Crab nebula. These observations revealed that the X-ray source has a diameter of about 1 light year; the visible part of the nebula has a diameter of about 6 light years.

Studies of supernovae in other galaxies as well as in our own, make it clear that there are two distinct kinds of supernovae. One reaches an absolute magnitude of about —13 at its peak; the other, which has been called an *ultrasupernova*, reaches an absolute magnitude of —16. The manner in which these two kinds of supernovae fade after reaching their peak can be used to distinguish them. The ultrasupernova fades more regularly than the less luminous supernova. Since the luminosity at peak is consistent within each type, supernovae are used as distance indicators whenever they are observed.

Stars with Expanding Shells

There are stars that have spectra similar to the spectra of novae but that do not vary much in brightness. One such group is called the *P Cygni* stars (Figure 11-15). These stars have broad emission bands with absorption lines on their violet edges, indicating that each of them, too, is surrounded by an expanding shell of gas. The absorption lines of P Cygni itself give a velocity of expansion of about 70 miles per second, considerably less than that of a nova. The

P Cygni stars seem to be continually losing material at a rate that would amount to the mass of the sun in 100,000 years.

The *Wolf-Rayet* stars are similar to P Cygni stars in that they have broad emission bands with absorption lines on their violet sides, evidence of an expanding shell of gas about each star. The displacement of the absorption lines in a Wolf-Rayet spectrum indicates a velocity of expansion for the shell of close to 2,000 miles per second. The lines of ionized helium in their spectra show that these stars are hotter than the P Cygni stars; with temperatures in excess of 50,000°K they are among the hottest objects in the sky.

Planetary Nebulae

There are stars in the sky that have very obvious shells of gas about them (Figure 11-16). When these were first observed visually with smaller telescopes they gave the appearance of planetary disks and consequently were named *planetary nebulae.* (To be sure, this is a misnomer, for they have scarcely anything in common with the planets that revolve about the sun.) The gas in the shell is greenish in color because of several strong emission lines in the green and the lack of a continuous spectrum. The shells are observed to expand at the rate of 6 to 30 miles per second, considerably less than the expansion of the shells about even the P Cygni stars. The diameters of planetary nebulae range from ⅓ to 3 light years.

The gas in the shell is so tenuous that it would be considered a good vacuum on the Earth, yet we see the shell. Why? The overall dimensions of the planetary nebula in Aquarius (NGC 7293) are about 2 by 2½ light years, and the shell is approximately ½ light year thick. Since the nebula looks like a smoke ring but is actually a shell, it is apparent that we cannot see the gases very well where they are only ½ light year thick. This is the "hole" in the middle. When we observe the edge of the nebula, however, we look through a layer of gas thick enough to be visible.

Photographing planetary nebulae with plates sensitive to blue and violet light discloses a central star that may be very faint, if not invisible, when viewed through a telescope directly with the eye. Since these stars are very hot (their temperatures range perhaps from 50,000° to 100,000°K), they emit most of their light in the blue and violet region where the eye is not very sensitive. There is much evidence to support the suggestion that the central stars are actually white dwarfs. This leads to the tentative conclusion that a planetary

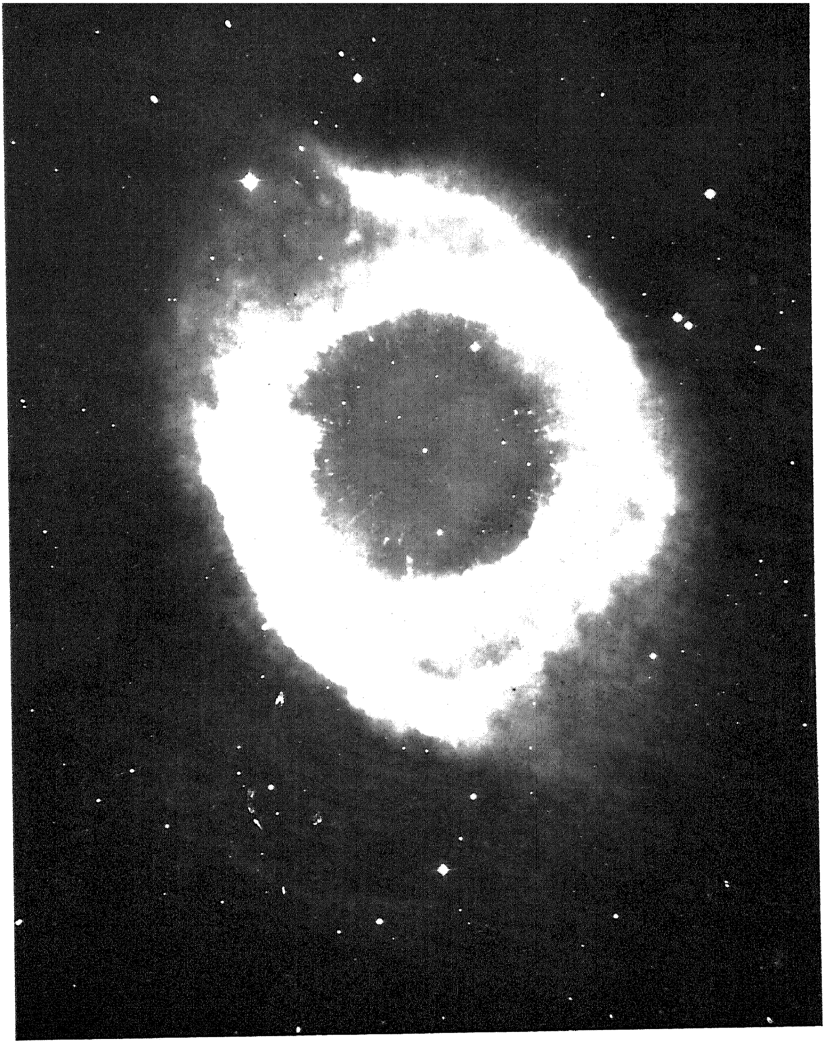


Figure 11-16 The planetary nebula NGC 7293. (Mount Wilson and Palomar Observatories)

nebula is a white dwarf in the making. Its shell eventually expands to invisibility and all that remains is the central white dwarf.

But what is a planetary nebula before the shell of gas forms? The 700 known planetaries are concentrated strongly toward the center of our galaxy and are not confined to the central plane of the Milky Way. Such a distribution is typical of Population II objects. Although this seems to eliminate all Population I stars as possible predecessors of planetary nebulae, it still leaves a choice between the many different kinds of Population II stars. Possible candidates for the job of supplying the universe with planetary nebulae are novae, Wolf-Rayet stars, RR Lyrae stars, the red-giant irregular variables, and perhaps a few more. But none of these seems to fill the bill completely; astronomers are left with another puzzle to solve. Because of the shell's slow rate of expansion, it is generally believed that planetary nebulae do not result from novae or Wolf-Rayet stars. Nevertheless, since they have undoubtedly ejected some material from their surfaces, they are related to the nova-shell star class.

Dwarf Novae

Not all nonstable stars are single stars. Nonstability can extend to double stars as well. An example of this is a group of stars which has been called the *U Geminorum* stars. It has been recently shown, however, that this class of stars is best grouped with another class, the *recurrent novae;* together, the two groups constitute a class better known as *dwarf novae.*

In 1943, A. H. Joy at the Mount Wilson Observatory discovered that the spectral lines of the dwarf nova SS Cygni shift periodically in a manner that leads to the conclusion that this star is really two stars very close together, revolving about a common center of mass. The spectra of both stars are observable, and Joy concluded that one star is a G5 dwarf and the other a B*e* star (the small *e* indicates that bright (emission) lines appear in its spectrum). Their period of revolution is only 0.276 day (about 6⅔ hours).

In 1954, M. F. Walker at the Lick Observatory discovered that the typical nova DQ Herculis is also a double-star system. The consideration of these two discoveries led R. P. Kraft to launch into a long-term program of study of this type of star. The results of this and other studies are more appropriately considered in Chapter 13, but it should be made clear that these two stars exchange gases. Gases from the larger cooler star pour over onto the smaller hotter star. This exchange of gases upsets the equilibrium of each star in such a way that flare-ups, eruptions, or explosions, occur from time to time. Walker has also observed very rapid changes in the brightnesses of this class of stars. Changes in the brightness of SU Ursae Majoris of as much as 0.1 magnitude have occurred in only 10 seconds of time. In addition to this rapid *flickering*, as it is called, the star

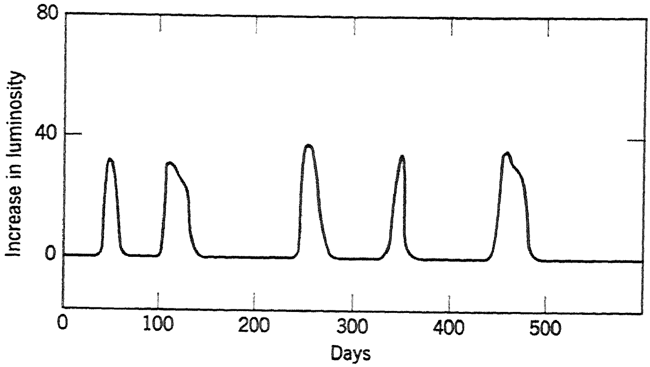


Figure 11-17 A schematic light curve for a U Geminorum star.

has other variations in brightness; it has been known to double its brightness within one hour.

But it is major changes in brightness which permit classification of these stars as dwarf novae. Their brightness may increase by a factor of 40 (Figure 11-17). These major outbursts do not occur with the clocklike regularity of Cepheids, but each star can be assigned an average interval of time between flare-ups. In general, the longer the interval of time, the greater the outburst. Stars with an average interval between outbursts of 10 days increase in brightness by a factor of about 8, and those that erupt about every 100 days increase in brightness by as much as 40 times. Between these major outbursts of energy, the dwarf novae evidence all of these short-range variations in brightness.

Exactly what causes the major eruptions, the variations of brightness within an hour, and the very rapid flickering is simply not understood. These stars present extremely difficult problems to the astronomer, yet the enticement of an ultimate solution and the realization that this solution will add to our understanding of the life processes of all stars urge the astronomer on. More observations and more theoretical work are required.

When all the objects discussed in this chapter are placed on the H-R diagram we have the apparently confused pattern in Figure 11-5. The diagram now looks considerably more complicated than it did in Figure 10-7, but the resulting difficulties are not insurmountable; they are only hurdles which we must jump in order to obtain a better understanding of the universe.

Basic Vocabulary for Subsequent Reading

Cepheid
Dwarf nova
Light curve
Nova
Planetary nebula
Population types
RR Lyrae star
Shell-type star
Synchrotron radiation
Velocity curve
Thermal radiation

For Further Reading

Struve, Otto, and V. Zebergs, *Astronomy in the 20th Century,* Chapters XV, XVI, Crowell Collier and Macmillan, New York, 1962.
Kraft, Robert P., "Pulsating Stars and Cosmic Distances," *Scientific American,* p. 48, July 1959.
Kraft, Robert P., "Exploding Stars," *Scientific American,* p. 54, April 1962.
Liller, William and Martha, "Planetary Nebulae," *Scientific American,* p. 60, April 1963.
Mumford, George S., "The Dwarf Novae," *Sky and Telescope,* part I, p. 71, February 1962; part II, p. 135, March 1962; part III, p. 190, October 1963.
Oort, Jan H., "The Crab Nebula," *Scientific American,* p. 52, March 1957.
Struve, Otto, "The Pulsating Star RR Lyrae," *Sky and Telescope,* p. 311, June 1962.
"Slow Fade," *Scientific American,* p. 86, September 1964.
"Expansion of Planetary Nebulae," *Sky and Telescope,* p. 316, June 1962.
"Expansion of Supernovae Remnants," *Sky and Telescope,* p. 209, April 1966.

Photographs of three Herbig-Haro objects, which are presumably gases in the process of forming stars. (Courtesy of George Herbig, Lick Observatory)

Chapter 12 *The Evolution of Stars*

The study of the evolution of stars is based largely on studies of the generation of energy in stellar interiors. Thus, since we can see only the light that is emitted by the surface, evolutionary studies must be largely theoretical.

Observation of the stars in their present condition (actually the condition they were in when they emitted the light that only now reaches us on Earth) tells us something of their structure and characteristics. To this information the astronomer must apply his general knowledge of matter and energy to deduce information about stellar interiors: their temperature, density, pressure, and chemical composition. Then he must calculate how a star of given characteristics would theoretically react to the conversion of nuclear energy into heat and radiant energy. He must also try to find some way to account for stellar formation.

When he has built up what looks like an acceptable hypothesis on the evolution of stars, he must check it with more observations. No doubt there are some predictions that can be made from his hypothesis; if they can be shown to be correct, the validity of the hypothesis becomes that much more likely. On the other hand, observation may prove the hypothesis incorrect and the astronomer must build a new one. But, having learned by each step in the procedure, he never again has to go back to the beginning.

The hypotheses that have been most successful in explaining the stars have all been based on the assumption that stars are gaseous throughout. In order to obtain a better understanding of the stars we should therefore investigate the laws that explain the behavior of gases.

Gas Laws

In studying a gas we consider three main factors: *temperature, pressure,* and *volume.* These three factors are related in such a way that if one changes, one or both of the others must change. Any one of the three may remain constant, but two will not remain constant while the third changes.*

The temperature of a gas is dependent upon the kinetic energy (energy of random motion) of the atoms composing the gas. For a given gas this means that an increase in the velocity of the atoms manifests itself as increased temperature.

The pressure of a gas is dependent on the momentum (mass multiplied by velocity) of the atoms. This means that if the velocity of the atoms in a constant volume of gas is increased, their momentum increases, and consequently the pressure of the gas increases.

The volume of a gas depends on the internal pressure of the gas as well as on the pressure confining it. For a gas in a balloon or in a cylinder the confining pressure may be supplied by the rubber in the balloon or by the walls of the cylinder. In the stars it is supplied by the gravitational field of the star itself. An atom cannot escape from a star unless it reaches the escape velocity.

If the temperature of a gas is increased by heating, the velocity of the atoms increases. Consequently their momentum increases and thus also the pressure. If the confining pressure is less than this increased internal pressure the gas will expand like an inflated balloon placed in a warm room. Conversely, if a gas is cooled the pressure will decrease and the gas will contract, if allowed to do so, like a balloon placed in a refrigerator.

---

* The general gas law is given as: $PV = nRT$, where P is the pressure, V the volume, n the number of moles of gas (the amount of gas contained in a quantity whose mass in grams is equal to the molecular weight of the atoms of that gas), and T the temperature on the Kelvin scale. If the pressure is given in units of the Earth's atmosphere (1 atmos = 14.7 lb/in.^{2}) and V in liters, then $R = 0.082$.

On the other hand, if a gas is compressed to a smaller volume by the confining pressure (for example, by a piston in a cylinder) its temperature increases. If it is allowed to expand (blowing out breath rapidly as against breathing it out gradually) its temperature decreases.

Gas Laws Applied to Stars

A stable star is one that is neither expanding nor contracting. Such a star is said to be in equilibrium, for all the forces acting on it are in balance. A system of checks and balances is necessary to maintain such a stable equilibrium: the heat energy generated in the core of the star must be allowed to escape in some way or other as rapidly as it is generated. So we must concern ourselves with the sources of energy and the ways by which this energy, once generated, may escape from the star.

The main source of energy in stars in the thermonuclear reactions (the proton-proton reaction and the carbon cycle discussed on p. 210ff) that convert hydrogen into helium, releasing heat and radiant energy. The rates at which both these reactions proceed is dependent on the temperature: the higher the temperature, the faster each reaction takes place. The reaction rates also depend on the density of the gas, and again, the higher the density the faster the reaction takes place. If the gas is not hot enough, neither of the reactions will take place and no energy will be generated from this source.

There is a second source of energy in the stars, but it is largely potential, since it is used only when required. This is the *gravitational potential energy* of the gas itself. If a star were allowed to contract and thus compress the gases in its interior, the temperature of those gases would increase as gravitational energy would be converted into heat energy.

To maintain equilibrium, however, the energy generated in the interior must be allowed to escape. This may happen in one of two ways. First, the star could radiate electromagnetic energy (light, etc.) into space in the manner of all stars. But a star may also rid itself of some of its heat energy by expanding and converting heat into gravitational potential energy. In expanding, the gases cool and the star becomes larger. We might recall that the star is not in equilibrium while it is expanding or contracting.

It is the combination of these sources of heat energy and of outlets for the disposal of the energy that constitutes the system of checks and balances. The gravitational force acting on the gas acts to compress the star. If the star were allowed to contract, the temperature and density of the interior would increase, and an increase in the reaction rate of the thermonuclear reactions would result. The increased reaction rate would generate more heat and the temperature would increase even more, causing the star to expand to equilibrium size.

The internal pressure acts to force the star to expand. If the star were allowed to expand, the temperature and density of the interior would decrease, and a decrease in the reaction rate of the thermonuclear reactions would result. Consequently, less heat would be generated, the temperature and pressure would drop, and the star would contract back to equilibrium size.

Thus the star's gravitational potential energy and the reaction rate of its thermonuclear reactions are balanced. The gravitational potential absorbs heat if the star expands and releases heat if the star contracts. The reaction rate increases if the temperature increases and decreases if the temperature decreases. Practically all the energy generated in the stars is released from them in the form of radiant energy; the gravitational outlet acts only as a safety valve.

Energy Transport and Stellar Structure

A star's energy is generated deep in its interior where the temperature is highest; however, the energy must reach the surface before it can radiate into space. If a star were completely opaque to all forms of energy transport, the core would heat up to such an extent that the star would eventually have to explode to rid itself of the energy generated. On the other hand, if a star were very transparent, its energy would leak out faster than it is being generated by the thermonuclear reactions. The star would then have to contract to supply the necessary heat to maintain the temperature and pressure required to support the overlying layers. A stable star also adjusts itself so that the amount of energy generated equals the amount of energy radiated.

The energy generated in the interior is transported to the surface in three ways: *conduction*, *convection*, and *radiation*. Conduction is the familiar form of heat transport along a metal rod. If one end is placed in a fire the other end will heat up, because the atoms in the hot end transfer their momentum to those in the cooler end simply by knocking into and jostling them.

Heat transport by convection is the actual movement of the hot object or gases. A central heating system in a house transports heat by convection if the hot material (air, water, or steam) itself is moved from the furnace into the rooms.

The transport of heat by radiation is accomplished when a hot object emits electromagnetic waves. An example is the familiar bathroom electric heater without a fan. If a fan is attached to the heater, heat will be transported by convection as well as by radiation. Heat energy is transported from the sun to the Earth by radiation.

A star's center contains a core through which heat may or may not be transported by convection currents (Figure 9-3). If there are convection currents in the core, it is evident that the material in the core will be thoroughly mixed. Surrounding the *convective core* is the *radiative zone* where energy can be transported only by radiation, and surrounding all this is the *convective envelope*. The convective envelope is the outer layer of the star; from the outside surface of this convective layer the energy is radiated into space. The convective envelope must be rather homogeneous in composition, since convection currents act to mix the gases.

According to current theories, not all stars have convective cores and envelopes. Of the stars on the main sequence, the sun, a G2 star, seems to be of the latest spectral type to have a convective core. Main-sequence stars later than F0 have convective envelopes, and thus the sun has a convective envelope as well.

Once the thermonuclear reactions begin in the hot center of the star they convert hydrogen into helium, increasing the amount of helium in the convective core and depleting the amount of hydrogen. Since the reactions demand a supply of hydrogen, the energy is eventually generated on the surface of the growing helium core where the high core temperature meets the hydrogen.

The radiative zone does not permit convection currents. Consequently the gases in this region are not well mixed, nor do they mix with the helium core. Since the temperature is so very high on the surface of the helium core, the energy is emitted in the form of X-rays that are successively absorbed and re-emitted by the atoms of the radiative zone as they proceed outward. The X-rays become longer in wavelength (and are then often called *softer* X-rays) as they are re-emitted, until they reach the convective envelope. Here they are absorbed by the gases, which consequently heat up, expand, and thus rise to the surface of the star, radiating their energy into space. In radiating their energy they cool and therefore sink back into the envelope to become heated again. The granulations on the surface of the sun (the only star whose surface we can see) are direct evidence of these convective currents in the envelope (Figure 9-4).

All the energy must be transported by radiation at least part of the way to the star's surface, even though convection may transport it over other parts of the distance. But radiation exerts a pressure and the intense X-ray radiation in the stellar interiors can exert a significant amount of pressure to help support the overlying layers of gas. The radiation pressure increases as the temperature of the core increases, and consequently becomes important in the very hot stars.

Energy transport by conduction is not effective in stellar interiors until the gases become highly compressed and completely ionized, and thus degenerate. Under this condition, heat conduction becomes the dominant form of energy transport, for the very fast-moving electrons conduct the heat from the hotter to the cooler regions of the degenerate gas. This is called *electron conduction* and is important only in the cores of red giants and throughout the interiors of white dwarfs.

Theory of Stellar Evolution

The process by which stars form from a gaseous nebula is not well known, but it has been speculated that a parcel of gas must by some method, perhaps by a whirling motion, detach itself from the rest of the nebula and commence the process of contraction. Such a contracting parcel of gas is called a *protostar,* for it is not hot enough to support thermonuclear reactions and thus does not emit visible radiation. As the gas continues to contract, gravitational potential energy is converted into heat energy and the temperature increases. When the central temperature approaches 500,000°K, a reaction involving deuterium (hydrogen atom with mass 2) may begin if the primeval gas contains deuterium. The deuterium would combine with a proton to produce helium 3, and this reaction would help increase the temperature of the protostar until the proton-proton reaction is initiated.

Soon after the proton-proton reaction begins, the protostar presumably becomes a full-fledged star and can be placed on the H-R diagram according to its luminosity and surface temperature. Since the central temperature and thus the luminosity depend on the mass of the star, the H-R diagram will separate new stars according to their mass. If the mass of the protostar is, let us say, 20 times that of the sun, it contains more gravitational potential energy to convert into heat than the sun. In fact, it *must* produce a hotter interior to create an internal pressure high enough to support the greater amount of gas.

For the more massive stars the contracting process is not stopped by the relatively slow proton-proton reaction. Contraction continues until the central temperature becomes high enough to start the carbon cycle, which proceeds more rapidly and is thus able to produce enough energy to support the gas above the core. The proton-proton reaction does not stop when the carbon cycle begins. Both reactions proceed simultaneously, but if the temperature becomes greater than about 25,000,000°K the carbon cycle begins to supply more energy. Presumably, both reactions take place in the sun, with the proton-proton reaction supplying more energy than the carbon cycle.

The protostar enters the H-R diagram on the right-hand edge (Figure 12-1) and proceeds up and to the left as both its temperature and luminosity increase. As soon as the nuclear reactions in the interior produce enough energy to stop the contraction by balancing the pressure there is a slight gravitational adjustment; the star drops slightly in luminosity and settles down on the main sequence. The line which represents the positions of stars newly arrived on the main sequence is called the *zero-age main sequence.* The shape of the curve representing the initial contraction stage for stars is apparently nearly the same for all stars regardless of their mass. A different mass only shifts the position of the curve up or down on the diagram.

Since a very massive star will have a high temperature and great luminosity, it will begin its life as a full-fledged star somewhere in the upper left-hand corner of the H-R diagram. A star of 16 solar masses will become a B0 star. A star of about $\frac{1}{3}$ solar mass will become a late M-type star. The stars of intervening mass will fall at positions between these two and form the bulk of the main se-

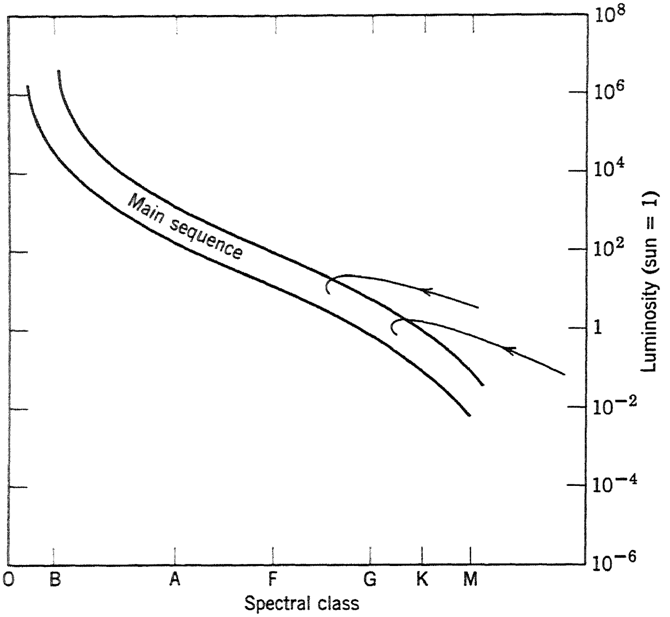


Figure 12-1 The evolutionary tracks of stars of moderate mass in the initial contracting stage.

quence. Thus, the main sequence represents stars that are in the first stage of their development as full-fledged stars. Stars more massive than 16 or 20 solar masses are rare and become the O-type stars.

Thus far, we have neglected possible variations in the chemical composition of the stars. We have considered them to be composed very largely of hydrogen. However, differences in chemical composition do occur, even in young stars. These consist of a difference in percentage of helium, or of some of the heavier elements. These differences raise or lower the star's path on the H-R diagram. The larger the percentage of hydrogen, the lower the curve will appear on the diagram. The possibility of differences in chemical composition makes us uncertain about the placement of the curve for any one star even though any differences that may exist would be slight, since hydrogen is by far the most abundant element.

A star spends a relatively short time in the protostar stage. A star the mass of the sun will contract from a portion of a nebula into a star on the main sequence in perhaps 50,000,000 years. As will be seen later, even such a span, long compared to a human life, is short in comparison with the entire life of a star. It is estimated that a star of 9 solar masses takes about 21 million years to contract and find a place on the zero-age main sequence. A very massive star, because of its stronger gravitational field, must contract even faster, perhaps in 500,000 years. A star with a 0.2 solar mass is estimated to take as long as 600 million years to contract from its initial stages to a full-fledged star. It is only because a star spends such a short time on this portion of its *evolutionary track* (its path on the H-R diagram as it evolves through various stages of its life) that we do not find many stars in this region of the diagram.

Once on the main sequence the star remains there for a considerable length of time, but not indefinitely. A star whose mass is equal to that of the sun may remain there for 10 billion years. A star with a mass less than that of the sun will remain on the main sequence longer. A very massive star will burn its hydrogen fuel up faster and consequently may remain only a few million years. The length of a star's stay on the main sequence depends on its mass; its evolutionary track as it leaves the main sequence also depends on its mass. The evolutionary tracks of stars of moderate mass are better known than those of very massive stars, and we will discuss them first.

When about 10% of the hydrogen fuel has been burned, the core begins to contract, releasing energy that forces the outer regions of the star to expand and hence cool. However, the increase in luminosity resulting from an increase in the size of the star exceeds the decrease in luminosity caused by the drop in surface temperature, so the star's net luminosity increases, moving it upward on the H-R diagram. The decrease in surface temperature simultaneously carries it to the right; its high luminosity and low surface temperature have made it a red giant (Figure 12-2). At a certain point, depending on its mass, it ceases to expand. In establishing this temporary equilibrium it may pulsate slightly, presumably becoming an irregular variable.

Oddly enough, the outer regions of the star in this stage are expanded and thus have a low temperature; at the same time the contraction of the core results in a higher central temperature and density. Because of this higher temperature and density in the core, some of the helium is converted into carbon and other heavier elements. These new nuclear reactions, of course, supply even more energy for the star, preventing it from contracting immediately.

How long the star remains a red giant is not well known, but

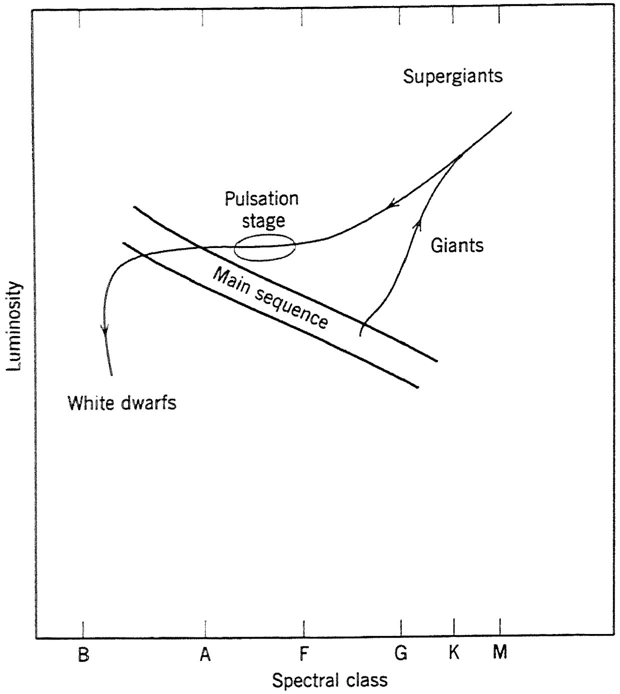


Figure 12-2 The evolutionary tracks of stars of moderate mass after leaving the main sequence.

it cannot retain this position as long as it does its main-sequence position. This follows from the fact that there are many more stars on the main sequence than anywhere else on the H-R diagram, which indicates that the average star spends a longer part of its life there than it spends in any other stage of its early evolutionary track.

The reasons why a star leaves the main sequence and becomes a red giant are known to the extent that astronomers feel they are at least (and at last) on the right track. Calculations by R. Härm and M. Schwarzschild at Princeton University for a star with a mass 1.3 times that of the sun indicate that, by the time the star reaches the red-giant stage, its helium core has grown to nearly 40% the mass of the star. By this time the core material has been subjected to very high pressures and temperatures. Under these extreme conditions of temperature and pressure, the core material is compressed into an exceedingly compact form of matter—*degenerate gas.* Degenerate gases (see p. 265) have exceedingly high densities and are not described by the same laws which describe the gases we have been discussing. The hydrogen gas overlying the core continues to "burn," but the radiative zone is not able to transport all of the energy generated by the core. As a result the core temperature increases. When it reaches 80,000,000°K (!), the helium starts to burn by converting into more massive elements. Because the core material is degenerate, it cannot yet expand; therefore, the temperature increases still more, and the helium-burning proceeds more rapidly. This rapid helium-burning is limited, however, for when the core reaches the fantastic temperature of 350,000,000°K it is forced to expand and the gases become nondegenerate.

The very rapid burning of helium before the core expands is called the *helium flash*, since it lasts for only a few thousand years. The expansion of the core causes the temperature to drop and the helium flash ends.

These same calculations indicate that the star's position on the H-R diagram does not remain constant during the helium flash. The star apparently moves from the red-giant stage slightly downward and to the left along the nearly horizontal giant branch of the H-R diagram. During part of this transition, instabilities arise and the star becomes a pulsating variable; which pulsating variable it becomes depends in part upon its mass. Presumably a more massive star becomes a long-period variable, a less massive star either a short-period variable or an RR Lyrae star. But the star's chemical composition may also be a factor in determining which type variable it becomes. There is theoretical evidence to suggest that the star may have a pass through the helium-flash stage more than once. Consequently, the star may move back and forth along the giant branch, becoming a variable more than once.

About the time the star crosses the main sequence, its surface temperature reaches a maximum and its luminosity decreases. The star sinks into the last stage of stellar evolution and becomes a white dwarf!

The general shape of the evolutionary track described is the same for all stars of moderate mass, even though the location of the track differs for stars of different masses. As the position of a star along the main sequence depends on its mass (the least massive star appearing at the bottom), so does the location of the star's evolutionary track. With increasing mass, the evolutionary tracks appear ever higher on the H-R diagram. A more massive star will be more luminous at any point along its track than a less massive one at a corresponding point. Furthermore, the more massive a star is, the faster it will evolve along its evolutionary track. The stages through which a star passes during its evolution are represented in part by the different types of stars already discussed. If one stage of a particular evolutionary track is populated by more stars than any other, we can assume that stars of that particular mass must remain in that stage of evolution longer than in any other stage.

OBSERVATIONAL EVIDENCE OF STELLAR EVOLUTION

The observational evidence for such a hypothesis of stellar evolution is quite strong. If stars form from gaseous nebulae, then we ought to examine nebulae in search for evidence of recent star formation. If in that search we observe the Rosette nebula we will find a large aggregate of gas with many very luminous stars in the center and with many dark globules of gas projected against the bright nebula (Figure 12-3). Continued search would reveal other nebulae that also exhibit these globules, which, it might be speculated, are protostars in the process of contracting. Their central temperatures are not yet high enough to support thermonuclear reactions; in fact the gas is still quite distended. They are simply blobs of gas and dust which may be contracting.

Before the Main Sequence

Evidence that these dark blobs of gas and dust might indeed be protostars is given by the *Herbig-Haro objects.* These are also found in nebular gases and each has a stellar or semistellar nucleus surrounded by a small emission nebula. Figure 12-4 shows a group of Herbig-Haro objects as they appeared in 1947, then again in 1954, and still again in 1959. The photographs clearly reveal new centers from which light is emitted. Presumably these are newly forming stars! Observations of their spectra indicate that these objects resemble in many ways another class of stars called the T Tauri stars.

In the constellations of Taurus and Orion there are other nebulae that contain *T Tauri stars*—a class of stars that are not very luminous, that have emission lines in their spectra, that are irregularly variable in light, and that are invariably associated with nebular material. All of the more than 500 T Tauri stars known are of a later spectral type than F8, and each is more luminous than a main-sequence star of the same spectral type. These two characteristics indicate that

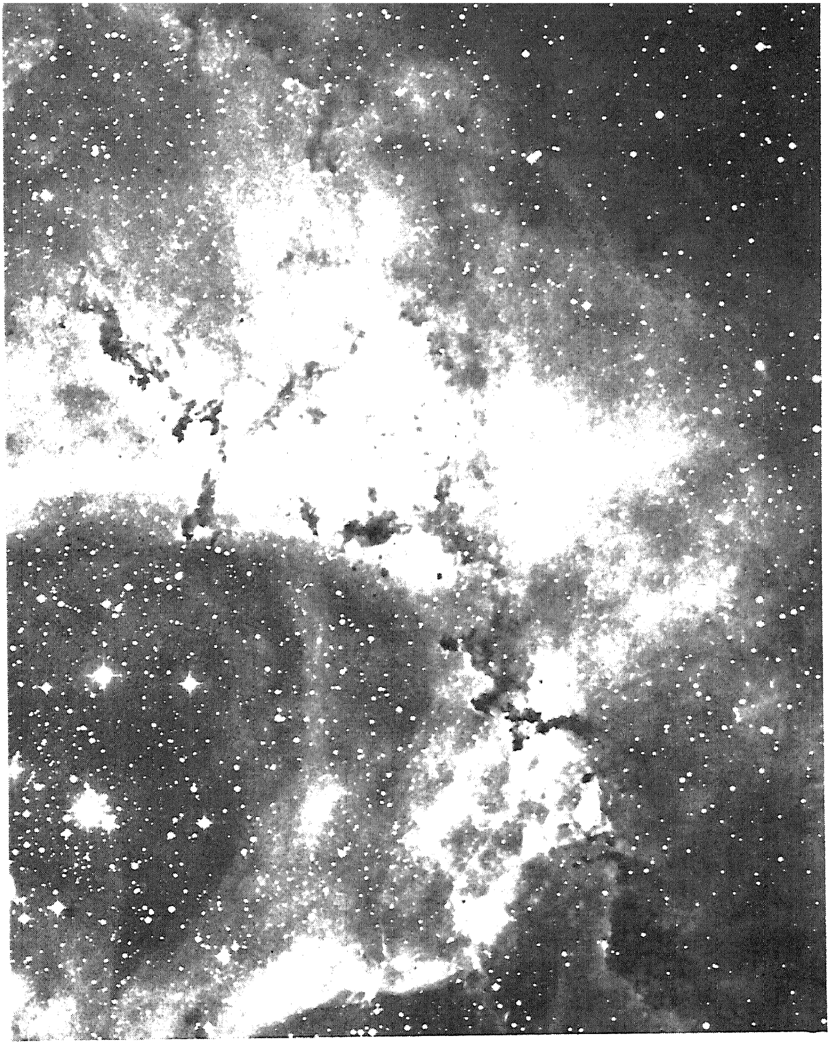


Figure 12-3 The Rosette nebula (NGC-2237) showing dark globules that may be protostars in an early stage of contraction. (Mount Wilson and Palomar Observatories)

they may be to the right of, as well as above, the main sequence. Such a position on the H-R diagram is predicted for a newly contracting star following an evolutionary track that leads to a point on the main sequence to the left of, and slightly below, its present position as a T Tauri star.

T Tauri stars have emission lines in their spectra which indicates

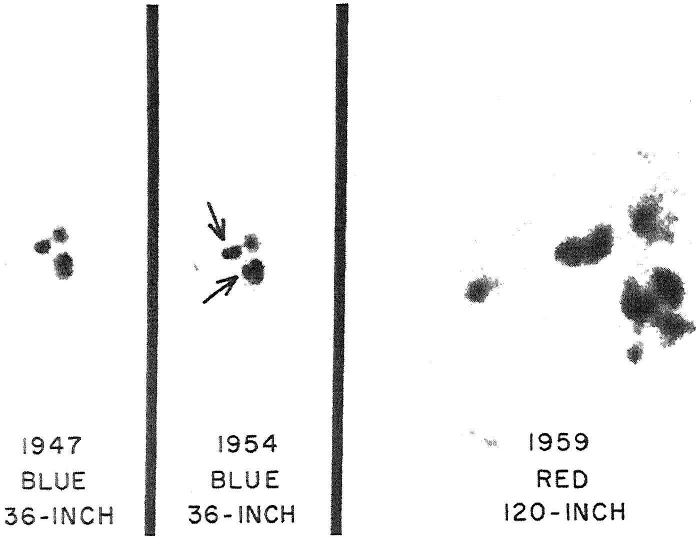


Figure 12-4 Herbig-Haro No. 2 in Orion taken in three different years; two in blue light (with a 36-in. telescope) and one in red light (by the 120-in. telescope). There has been further brightening of one of the new nuclei since 1959. (Courtesy George Herbig, Lick Observatory)

that they have either extensive atmospheres or nebulous material immediately surrounding them. The fact that they are invariably associated with nebular gases places them in what present theory considers the necessary environment for stars in the early stages of formation. However, they do move with respect to the clouds in which they are now immersed, which indicates that they will eventually leave these clouds. Since T Tauri stars are generally not found outside gaseous clouds, we must conclude that they lose their identifying characteristics before or when they leave the clouds. Thus, the stage of development represented by the T Tauri stars may be a short one, embracing perhaps only 10% of the time required for a star to contract onto the main sequence.

On the Main Sequence

Recent studies indicate that it may be possible to learn something about the length of time stars have spent on the main sequence. The time indicator is lithium. It is well known that the lithium nucleus is formed in one part of a star, but it cannot survive the high temperatures encountered in stellar interiors. It is also known that young stars such as T Tauri stars have a great deal more lithium in their atmospheres than has the sun. But George Herbig of the Lick Observatory has observed a range in the lithium content of 60 G-type stars in the solar neighborhood.

If the gases in the atmosphere mix to some extent with the gases deeper in the star, the lithium content in the atmosphere should decrease with time, for as the gases are mixed some of the lithium is carried into the deeper and hotter regions of the star where it is destroyed. At an age of 4.6 billion years the sun has very little lithium left. Younger G-type stars presumably have more. Before ages can be assigned, however, something must be known about the percentage of lithium in the atmosphere of the new star, the rate at which the atmospheric gases mix with the gases beneath the photosphere, the temperature under the photosphere, and the rate of destruction of lithium at those temperatures. Continued study of this problem promises to help our understanding of stellar evolution considerably.

After the Main Sequence

During their life on the main sequence, the stars burn about 10% of their hydrogen content; they deplete the hydrogen in their core. Evidence of what happens to the very massive stars after they have depleted this hydrogen is given by the *Beta Canis Majoris stars.* This is a group of stars composed of early B-type stars that are more luminous than their main-sequence counterparts (Figure 11-5), and that are variables in every sense of the word.

The brightness of a Beta Canis Majoris star is evidently related to its period; the stars of earlier types are more luminous and have longer periods than those of later types. The range in periods for the whole class is from 3½ to 6 hours. They also exhibit periodic fluctuations in radial velocity, indicating that they pulsate. One of the more remarkable characteristics of a Beta Canis Majoris star is that its period increases in a regular fashion. Even though this increase is small, less than 3 seconds a century, it is nonetheless observable. Because the rate of pulsation is so rapid the star goes through as many as 2,000 oscillations each year, which enables us to determine its period to an accuracy of 0.03 second. Since at least one Beta Canis Majoris star has been observed for 30 years, a change of 3 seconds in a century or 1 second in 33 years is relatively easy to measure.

The late Otto Struve suggested that the change of period of a Beta Canis Majoris star may actually be interpreted as an evolutionary

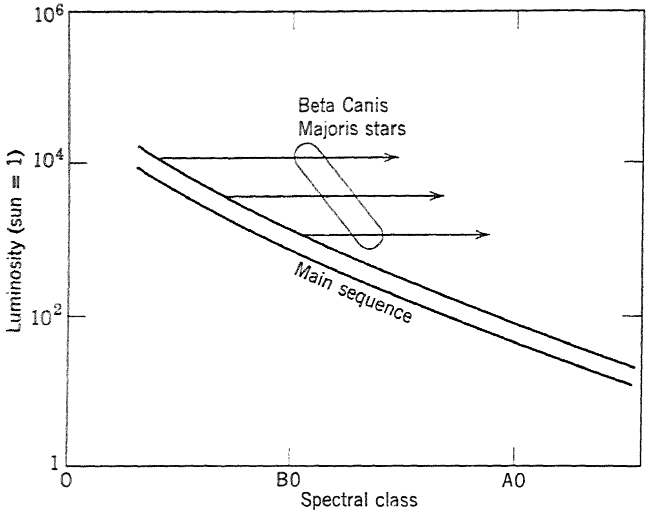


Figure 12-5 **The Beta Canis Majoris stars may represent a stage in the evolution of massive stars.**

change. Struve proceeded from the assumption that a Beta Canis Majoris star once existed on the main sequence to the left of where it now resides (Figure 12-5). As its hydrogen was converted into helium, it expanded; thus, its surface temperature decreased, taking it to the right of its original position on the main sequence. Pulsation then set in when it entered the region defined by the Beta Canis Majoris stars. As the expansion continued, the density of the star decreased and the length of the pulsation period increased correspondingly.

If this interpretation is correct, this is, among the stars, the only evolutionary change that can now be followed and studied in some detail. It is not surprising to find that our first observation of what we suspect to be an evolutionary change occurs in the very massive stars, for these are the stars that evolve the most rapidly.

If the very massive stars on the main sequence are less than 1,000,000 years old, they must have been born only recently on the astronomical time scale. (The Earth is about 4.6 billion years old.)

As we see in Chapter 13, stars are frequently found in large groups called *clusters*, each held together by the gravitational field set up by the many stars. Since individual stars, like the sun, are so very far apart, the gravitational effect of any one on another is almost

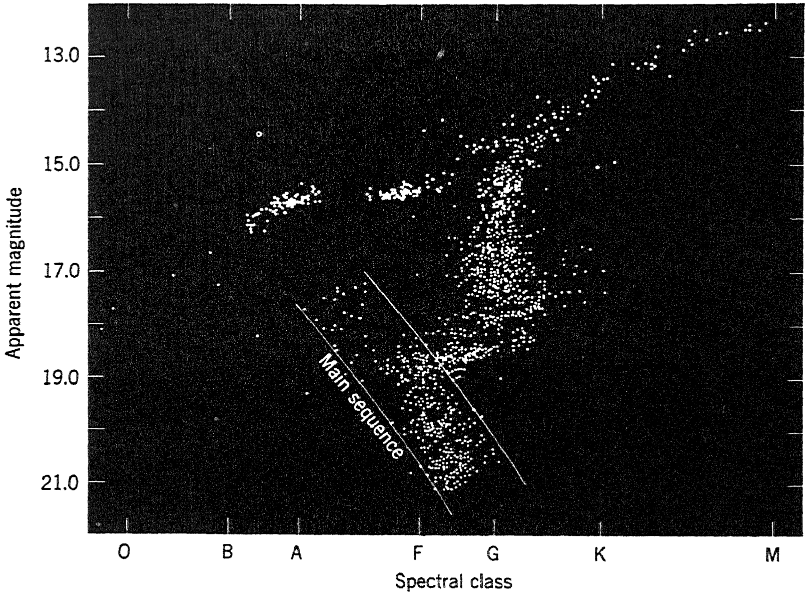


Figure 12-6 An H-R diagram of the globular cluster M3 according to H. C. Arp, W. Baum, and A. Sandage.

negligible. Consequently, clusters could hardly have been formed after the individual stars, separated by great distances, were formed. It is therefore generally assumed that all the stars in a cluster were formed at the same or nearly the same time and in a region of space roughly the same size as the cluster now occupies.

If every star does go through an evolutionary process the rate of which depends on its mass, in any one cluster we would expect to find the more massive stars in a more advanced stage of evolution than the less massive ones. That this is actually true is demonstrated by the H-R diagram (Figure 12-6) of a globular cluster (see compact cluster in Figure 13-14). The stars that have already strayed from the main sequence in their evolutionary track are the more massive stars; those still left on the main sequence are less massive. That point on the main sequence where the stars branch off toward the red-giant stage is called the *turn-off point*. It should be noted that very few stars remain on the main sequence above the turn-off point; however, the fact that some stars are on the main sequence above the turn-off point leads us to suspect that these stars may have been formed later than the bulk of stars in the cluster. Therefore, not all the stars in the cluster are of precisely the same age. This diagram, then, does not represent the evolutionary track of any one star of given mass, but identifies points on evolutionary tracks of stars of differing masses whose ages are nearly the same. The stars that are still on the main sequence will eventually evolve into red giants, but these less massive stars will presumably be less luminous than the red giants appearing in the cluster at the present time.

The gap in the upper branch of the diagram is of particular interest; it is here that the RR Lyrae stars in the cluster would be located, were they included on this diagram. According to current theory, the stars pass through the RR Lyrae gap (or Cepheid gap) during their helium flash when the stars move horizontally along the giant branch. It is of great interest to realize that a well known Cepheid, RU Cameloparadalis, has become an ex-variable star! The variability of this star was first recognized in 1907 and since that time its period has been established as just over 22 days, with a range in apparent magnitude of 8.2 to 9.1. In 1961 and 1962 its variations were normal, but in 1963 the extent of its variations decreased. Its light variations continued to decrease in 1964 and finally in 1965 they essentially stopped. After that time only slight variations (less than 0.04 magnitude) have taken place and these are irregular. This is the first observation of a cessation of a regular variable, and it has been confirmed at a number of observatories. Certainly this star will be watched closely for a number of years.

The evolutionary tracks of the very massive stars are not at all well known. There is some indication that they do evolve to the right of the main sequence and are similar in at least this respect to the less massive stars. But where they go from there is a mystery. One suspicion is that they do not evolve very far to the right of the main sequence. Evidence for this is offered by the fact that there are no super red giants which would correspond to the evolution of an O-type star that is evolving along a track *similar in shape* to the evolutionary tracks of the moderately massive stars. When a moderately massive star evolves off the main sequence it becomes more luminous. The most luminous stars in the universe (except novae), however, are the O-type stars and not the red giants. Accordingly, the O-type stars do not become more luminous when they leave the main sequence.

The situation is made even more difficult by the fact that there is no red giant whose mass we know with any degree of certainty. Consequently, we have only indirect evidence of what type of star a red giant was when it was on the main sequence.

If the stars leave the main sequence by expanding into red giants, it is not yet entirely clear just what happens to them immediately after the red-giant stage. Does the helium-flash stage apply to stars of all masses? In which direction do stars move along the horizontal giant branch? Do stars move back and forth along the horizontal giant branch? These and other questions have yet to be answered.

It is clear, however, that most stars eject vast quantities of gas back out into space. The sun has a solar wind which presumably is typical of all stars on the main sequence. Observational evidence makes it clear that both the blue giants and the red giants eject gas in the manner of a "stellar wind," but eject much more gas than the sun. Furthermore, it is clear that at some time near the completion of the helium flash, many stars eject matter more violently—recall the P Cygni stars, the Wolf-Rayet stars, the planetary nebulae, and the supernovae. There is evidence to support the idea that as the degenerate helium core builds up and its temperature increases, the increased radiation pressure literally blows gases away from the star. A recent study of the central stars of 65 old planetary nebulae by G. O. Abell of the University of California at Los Angeles indicates that these stars are just entering the white-dwarf stage. They have an average surface temperature of 50,000°K and an absolute magnitude of +5, placing them in the white-dwarf region of the H-R diagram. Their average radius is estimated to be about 0.1 the radius of the sun. It appears that eventually all of the gases overlying the degenerate core are blown away, and the core becomes the white dwarf.

White Dwarfs

Once a star has become largely degenerate we recognize it as a white dwarf. Because white dwarfs are composed mostly of degenerate matter, they do not obey the laws that explain the actions and structure of other stars. The laws explaining the white dwarfs were derived to a large extent by S. Chandrasekhar of the Yerkes Observatory, University of Chicago.

As we noted on p. 265, atoms no longer exist in degenerate matter—the nuclei and electrons become closely packed together. This limits the amount of space available to each particle, and its motion is consequently reduced. Some of the electrons manage to reach velocities that approach the velocity of light, but many are bound to a very small volume of space within the star and are not free to move at ramdom. Since the nuclei are more massive, they are not rigidly bound and, although still restrained, are freer to move. If any protons (hydrogen nuclei) remain, they are squeezed out of the central regions of the degenerate gas by the more massive nuclei and float on the surface. The electrons remain inside the star because the electric field set up by the nuclei attracts them.

One of the seemingly strange characteristics of degenerate gas is often called the *Chandrasekhar's limit.* Chandrasekhar pointed out that the size of a degenerate core or white dwarf depends almost entirely upon the amount of matter it contains, but that if the *mass is increased the diameter will decrease.* An increase in the amount of matter increases the gravitational field and simply squeezes everything closer together. Consequently, as the mass of the degenerate core increases by converting hydrogen to helium on its surface, it shrinks; but there is a limit! Chandrasekhar showed that a degenerate helium core cannot exceed a mass equal to 1.2 times the solar mass. With this mass it would shrink down to a diameter of zero! It is supposed that as the core approaches this limit, its temperature reaches such a high value that the core sheds its overlying layers of nondegenerate gas. If the star begins its life on the main sequence as a B-star, its mass is about 16 times that of the sun. During its succeeding evolution it ejects gases, first as a main-sequence star, then as a giant, then (perhaps) more dramatically, to leave exposed only the naked core of degenerate material. That core must have a mass less than 1.2 solar masses, yet be more nearly the size of our moon than the sun.

Such a high concentration of mass in a small volume gives rise to gravitational forces at the surface that exceed the gravitational force on the Earth by as much as 3,000,000 times. Thus, a man weighing 180 lb on the Earth would weigh nearly 270,000 tons on the surface of a white dwarf and, needless to say, would not survive even if the temperature were bearable.

This tremendous force of gravity affects the surface of the white dwarf. The theory of white dwarfs indicates that the degenerate gas comprising most of the star ends rather abruptly and is surrounded by an envelope of nondegenerate gas about 65 miles thick. Above this there is an atmosphere of sorts, but it is not more than a few hundred *feet* thick! This atmosphere probably contains any residual hydrogen left over from the star's earlier stages, which not only converted most of the hydrogen into helium but also, as a last resort, ejected the remaining hydrogen, perhaps as a shell-type star.

The pressure inside a white dwarf no longer follows the same laws that prevail in other stars and is called *degenerate pressure* rather than *gas pressure.* Unlike gas pressure, degenerate pressure is unaffected by temperature and therefore does not permit the star to absorb heat energy by expansion. Thus, if local heating occurs in one region of the star, any nuclear reactions occurring would have to increase in rate, and the star would explode because of the increased release in energy. According to the theory of degenerate matter, moderate expansion to allow the star to absorb this excess energy by transforming it into gravitational potential energy is not permitted. On the other hand, both theory and observation indicate that in reality white dwarfs do not explode; it is presumed, therefore, that they have used up all their available nuclear energy in transforming one element into another. When a star becomes a white dwarf it no longer generates energy; it is left to cool by radiating energy from its hot surface into cold space.

Observational evidence of the white dwarfs supports the theory of degenerate gases. The observed masses of white dwarfs are all under the theoretical limit. The larger white dwarfs have diameters of 10,000 miles; the smallest ones known can have a diameter not bigger than 1,000 miles. The observed surface temperatures range from 50,000°K to 4,000°K.

The spectra of white dwarfs differ markedly from those of other stars. In general, if any hydrogen lines are present they are weak, and indicate a small amount of residual hydrogen only. The white dwarfs with observed temperatures of less than 8,000°K have no hydrogen lines in their spectra, only a few metallic lines; at least six white dwarfs have no lines whatsoever in their spectra. The spectral lines that do appear must originate in the atmosphere; consequently when no spectral lines appear at all we assume that the star has no atmosphere.

One star, WZ Sagittae, seems to have entered the white-dwarf stage since observational astronomy began recording data. When this dwarf nova exploded in 1913 and again in 1946, its luminosity increased some 1,000 times. Its spectrum, except for some emission lines, now classifies it as a white dwarf. The emission lines are presumed to be caused by the ejected material that is still expanding, yet remains

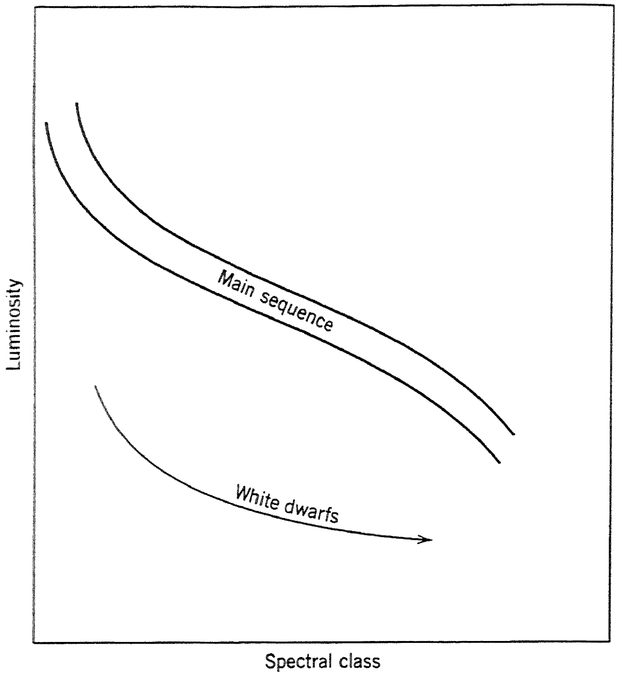


Figure 12-7 The evolutionary track of white dwarfs.

close enough to the star, and thus is hot enough to emit its own light.

According to theory and observation, WZ Sagittae has entered the last act of the drama of stellar evolution. Since it no longer generates energy, the white dwarf is left to cool by radiating its energy into space. As its surface temperature decreases, it radiates less energy. Consequently its temperature decreases more slowly, and it may require 3 billion years to cool to 7,000°K from its initial temperature of more than 12,000°K.

A star with a surface temperature of 7,000°K, however, no longer radiates white light—its light is a yellowish-white. Five billion years after this, its surface temperature will have dropped to 4,000°K and it will be red in color. The term "white dwarf" is retained, however, for this name has come to imply a star whose mass, size, and composition classify it as a ball of degenerate gas. The term "red dwarf" would imply that the star resides on the lower part of the main sequence (Figure 12-7).

Thus it takes about 8 billion years for a white dwarf to cool to a temperature of 4,000°K, and even then it is not dead. This time span amounts to nearly twice the age of the solar system and even so does not include the time the star has spent in the evolutionary stages which it has had to undergo before it became a white dwarf. Since white dwarfs with this surface temperature have been observed, we can conclude that our galaxy (in which they reside) is considerably older than the sun.

If stars spend such a long time in the white dwarf stage, then, following our supposition that the longer they remain in any given stage the more heavily populated that stage will be, there should be many more white dwarfs than any other type of star. But it appears that the white dwarfs comprise only 3% of the stars in our galaxy. This seeming inconsistency is eliminated, however, if we assume that our galaxy is no older than just the age required for the most massive stars formed early in its life to have evolved into such cool white dwarfs.

Nevertheless, a star whose temperature is 4,000°K still radiates energy, and heat will continue to flow from it until it reaches the temperature of space, since, according to the second law of thermodynamics, heat flows from a higher to a lower temperature. As the white dwarf cools it will do so ever more slowly. After billions of years its surface will have cooled to such an extent that it will radiate no energy in the visible region of the spectrum; it will become first an infrared white dwarf and, when it has cooled completely, a black white dwarf. At this stage the nuclei and electrons reach their lowest possible energy level and the white dwarf (now emitting no light) becomes essentially a single huge molecule. Since no more energy is available from either contraction or nuclear transformations, it must remain in this stage indefinitely, like a celestial tombstone.

Our present state of knowledge about stellar evolution is far from complete and will require a good deal more theoretical as well as observational work before we can determine the stage of evolution for any one particular star. In other words, for any single star we cannot yet tell accurately its age, its past, and its future existence. It is comforting, however, to find that some observations do fit in very well with present theoretical work.

In summary, stars form from very tenuous gases composed chiefly of hydrogen. They contract to the main sequence, begin converting hydrogen into helium, expand and eject matter back into space, and then pass through various stages before they finally contract into extremely compact balls of degenerate gas completely devoid of hydrogen.

Basic Vocabulary for Subsequent Reading

Beta Canis Majoris star	**Herbig-Haro object**
Evolutionary track	**T Tauri star**

For Further Reading

Baade, Walter, *Evolution of Stars and Galaxies,* Harvard University Press, Cambridge, Mass., 1963.

Ezer, D. and A. G. W. Cameron, "Early Solar Evolution," *Sky and Telescope,* p. 328, December 1962.

Greenstein, Jesse L., "Dying Stars," *Scientific American,* p. 46, January 1959.

Lovell, Sir Bernard, "Radio-emitting Flare Stars," *Scientific American,* p. 13, August 1964.

Sitterly, B. W., "Symposium on Abundances of Elements in Stars," *Sky and Telescope,* p. 11, January 1965.

Slee, O. B., et al., "Visual and Radio Observations of Flare Stars," *Sky* and *Telescope,* p. 83, February 1963.

Struve, Otto, "T Tauri Stars and Associated Nebulosities," *Sky and Telescope,* p. 197, October 1961.

Struve, Otto, "Variable Stars of Low Intrinsic Luminosity," *Sky and Telescope,* p. 67, August 1962.

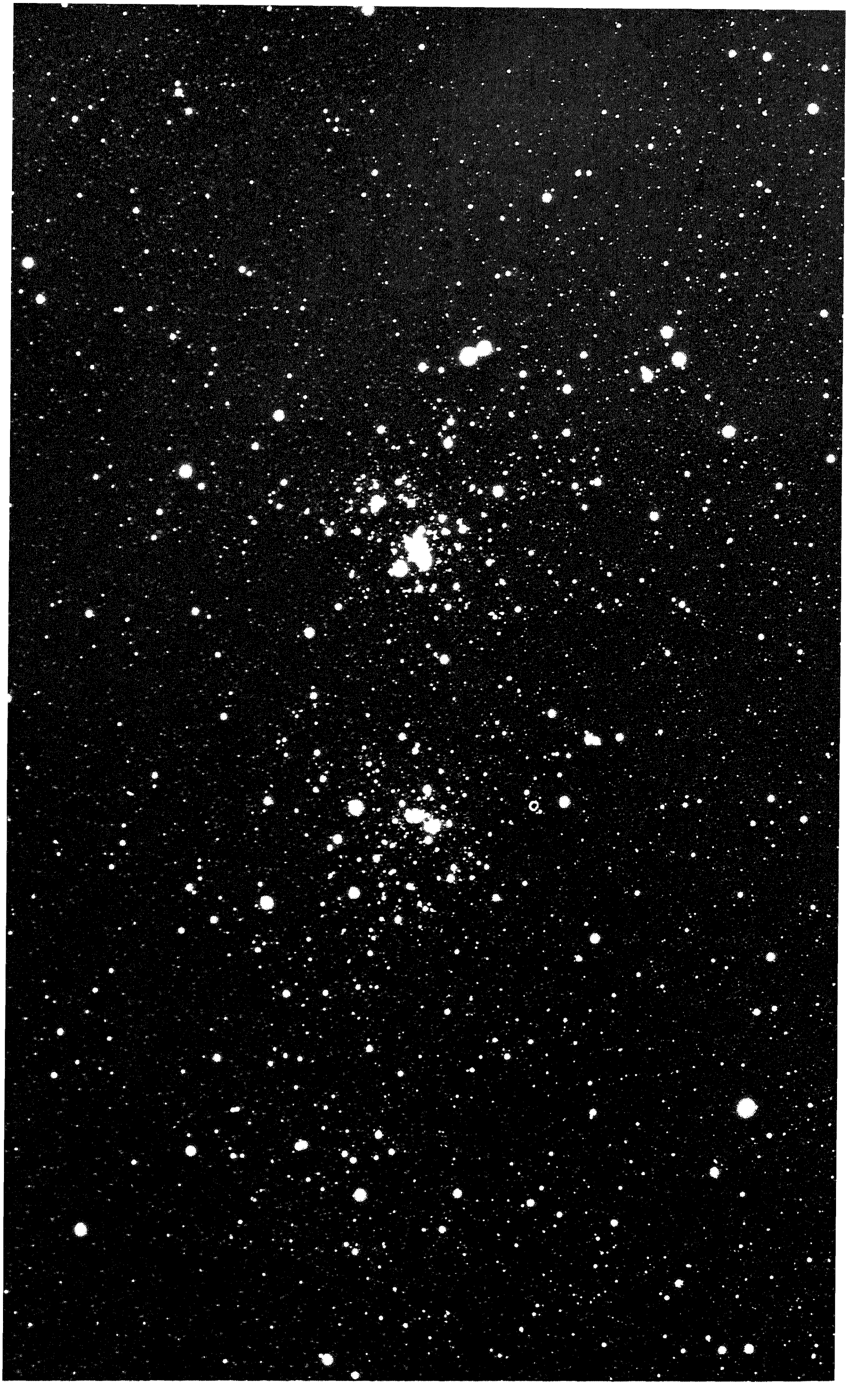
"Helium Flash in Giant Stars," *Sky and Telescope,* p. 275, May 1961.

"Infrared Glows Around Stars," *Sky and Telescope,* p. 158, March 1965.

"Lithium and Star Aging," *Sky and Telescope,* p. 82, August 1963.

"Smallest White Dwarf?" *Sky and Telescope,* p. 17, January 1964.

"An Ex-variable Star?" *Sky and Telescope,* p. 323, June 1966.



The double cluster, h and χ Persei. (Lick Observatory)

Chapter 13 *Multiple Stars and Clusters*

Our discussion of the evolution of stars has indicated that the stars are, over their lifetimes, very active and changeable, and that we have difficulty in observing their changes only because our life span is so short. This chapter will show that stars are also gregarious; single stars are actually less common than multiple stars. Some stars travel in pairs, others break the rule "two's company, three's a crowd," still others travel in fours, fives, sixes, tens, hundreds, and even thousands. All the stars in any one group are held together by their mutual gravitation. Some of these groups form very stable configurations and others are breaking up because the gravitational forces are not strong enough to hold the stars together. In a loosely held group each star goes its own way, parting company with its companions of travel through space.

Consider, for the present, only the *binary stars,* those that travel in pairs. We can learn a great deal more about them than we can about single stars. Some of the stars that form a binary system are so close together they seem to be in contact with one another and cannot be resolved into separate stars by even the largest telescopes. Others are so far apart, they seem almost like single stars until their motions through space demonstrate that they are traveling together.

Most binaries in our galaxy, however, do not present either of these extremes. The most common separation of binary stars is somewhat less than 20 A.U.—about the distance of the Jovian planets from the sun! Perhaps if the nebular disk around the sun had been different in some as yet unknown respect, Jupiter or Saturn might have become stars, and the sun would have had quite different companions. It has been suggested that all the stars on the main sequence may be accompanied by planets.

Visual Binaries

The binary stars that can be seen as two stars when viewed with a telescope are called *visual binaries.* Most of these need the resolving power of a telescope to separate them, but there are a few in which each member can be seen with the naked eye. Mizar and Alcor, at the bend in the handle of the Big Dipper, appear to the naked eye to be a binary system, but when viewed with even a small telescope Mizar is seen as two stars and Alcor is a third in the same field of view. With only a modest telescope many other stars that seem to be single when viewed with the naked eye are revealed to be double; Polaris is another example.

Some stars may appear to be visual binaries when in reality they are separated in space by many light years. Their near coincidence on the celestial sphere results from their being nearly lined up when viewed from the Earth. Such apparent binaries, called *optical doubles,* are fairly uncommon. The difference between an optical double and a visual binary can be noted by following the motions of the two stars. A visual binary forms a physical system, in that the stars are under one another's gravitational attraction. Therefore they travel through space together as they revolve around each other, or more precisely, as each revolves around the common center of mass of the system. Stars forming an optical double, on the other hand, usually have different space motions and do not comprise a physical system.

Since the force of gravity holding two stars into a physical system depends on their masses, we have an approach to the problem of measuring stellar masses. Fortunately this approach is not completely new to our discussion. The laws that explain the motions of two stars about their center of mass are the same laws that explain the motions of the Earth and moon about their common center of mass and the motions of the planets about the sun. We can use Kepler's three laws.

The first two laws can be restated as they were given on p. 97: the orbits of each star about the center of mass must be elliptical, and the line joining the star with that center of mass (the radius vector) must sweep out equal areas in equal intervals of time. The third law, however, must be given in a more complete form. As stated on p. 99 it does not take into account the masses of the sun and planets. Those masses could be neglected because the sun is so much more massive than any of the planets. But the stars in a binary system are likely to have more nearly the same mass. Consequently, it is Kepler's third law that gives us information about the masses of the two stars.

Kepler's third law was originally given as a direct proportion between the squares of the periods and the cubes of the orbital radii of any two planets. The more complete law is also given as a proportion, but rather than comparing any two planets it compares any binary system with the sun-Earth system.

$$\frac{M + m}{M_s + m_e} = \frac{s^3/P^2}{s_e^3/P_e^2}$$

Fortunately this equation can be greatly simplified by the proper choice of units of measurement. Since the separation of the stars in most binary systems is similar to the distance of the planets from the sun, we can measure these distances in astronomical units. Thus the radius of the Earth's orbit s_e becomes equal to 1. We can also select the year for the unit of time and by so doing make the period of revolution of the Earth P_e equal to 1. This enables us to eliminate two of the factors from the equation.

The sun's mass, 332,958 times that of the Earth, almost completely overwhelms the Earth's mass; there is no appreciable difference between the sun's mass $M_s = 332{,}958$, and the combined sun-Earth mass $M_s + M_e = 332{,}959$. Therefore we may neglect the Earth's mass and not feel that by doing so we will have damaged the validity of our calculations.

Lastly, we can measure stellar masses in terms of the sun. Thus the mass of the sun $M_\odot$ equals 1.

The formula now takes on a simpler form:

$$M + m = \frac{s^3}{P^2}$$

where M is the mass of one star (usually the more massive of the two) and m is the mass of the other, both in solar masses, P is the period of revolution of one star about the other in years, and s is their mean separation (the average orbital radius of one star about the other) in astronomical units.

In using this formula it is convenient to assume that the *primary* (the brighter star) does not move, and to establish by observation the orbit of the *secondary* (the fainter star) with respect to the primary component. In doing this we neglect the center of mass of the system, but Kepler's laws still hold for this *relative orbit.*

Assuming for the time being that the mean separation s and the period P can be measured, we are left with the sum of the masses and not with each individual mass. If the sum of the masses of a given binary system is found to be 10 solar masses, for example, we do not know whether the star's individual masses are 1 and 9, 5 and 5, or any other combination that adds up to 10.

We therefore need to establish some other relationship which when combined with the sum of the masses will yield each individual mass. This can be done only if we consider the motion of each star around the center of mass. The average distance of each star from the center of mass is equal to its mean orbital radius and depends on the individual mass of each star. The larger the mass the smaller the mean radius. This can be expressed as an inverse proportion:

$$\frac{M}{m} = \frac{r}{R}$$

where M and m are the masses of the two stars, M being the larger, and R and r are the mean orbital radii, R referring to the more massive star.

From this expression we see that the mass of the larger star is equal to the mass of the smaller multiplied by the ratio of their radii.

$$M = \frac{r}{R}m$$

Since the mean orbital radius of the less massive star r must be larger than that of the more massive star R, the quotient r/R must be greater than 1.

To see how this relation fits in with the one relating the sum of the masses, let us assume that the sum of the masses is equal to 10. Furthermore let us assume that r is four times larger than R and therefore r/R is equal to 4. We then have the following:

$$M + m = 10$$

$$M = 4m$$

If we substitute the value $4m$ from the second equation for its equivalent M in the first equation we have

$$4m + m = 10$$

Thus $5m$ equals 10, and m must equal 2. If m is equal to 2, then M must equal 8, and we have found the individual mass of each star in terms of the sun's mass.

A mathematician will say that if we have two unknowns we can only find the value of each if we have two equations. This we did. Our two unknowns were the individual masses of the stars, and the two equations were the expressions involving the sum of the masses and the ratio of the masses. Thus the mathematics involved are not too impenetrable, but the same cannot be said for the observations that must be performed to supply the data needed for the calculations.

We have three quantities that must be observed: the period of the secondary component about the primary, the mean separation of the two stars in astronomical units, and the ratio of the mean orbital radii of the two stars about the center of mass of the system.

The observational part of the problem is complicated because the orbits of the visual binaries are oriented more or less at random in the galaxy. Thus we see them at different angles. Some we see "face-on," that is, the plane of the orbit makes a right angle with the line of sight, and we can get a true picture of the shape of the orbit of the secondary about the primary. Most of the visual binaries, however, are inclined from this most desirable situation, and the shape of their orbits becomes distorted by foreshortening. If we look at the top of a round table situated at a short distance from where we are, its outline will be elliptical rather than circular (unless we are directly above the table). Elliptical orbits seen at an angle will also appear to change shape, although they will still appear as ellipses.

At the other extreme some of the orbits may be seen "edge-on," and the stars seem to move in straight lines back and forth in front of and behind each other.

The plane of the orbits of a binary system that makes a right angle with the line of sight must be parallel to the *plane of the sky*. The plane of the sky is that portion of the celestial sphere in the region of the orbit. To illustrate, we may place a small ring flat on the inside surface of a large globe with the observer at the center. The ring touches the globe at all points and thus the two are parallel.

The angle between the plane of the orbit and the plane of the sky is called the *inclination*. When the two planes are parallel the inclination is zero. If the plane of the orbit is tipped from the plane of the sky, the inclination is greater than zero. Such an orbit resembles the ring on the inner surface of the globe when one side of the ring is raised from the surface. The true shapes of the orbits are not seen when they are so oriented in the sky, and the astronomer must deduce the *true orbit* from its *apparent orbit* (the projection of the true orbit onto the celestial sphere). Fortunately, Kepler's laws hold true even for the apparent orbit. When the plane of the orbit is seen "edge-on" it makes a right angle with the plane of the sky and its inclination is 90°. A ring placed "upright" against the inner surface of the globe is in a comparable position.

To relate these observations to the mathematics already discussed let us use a simple example—when the inclination of the orbit is zero. If we observe the orbit of the secondary star about the primary, we obtain an ellipse of the same shape as each of the individual orbits about the center of gravity, but of different size. Kepler's laws again apply. The period of the secondary component about the primary component is the same as the period of each about their common center of mass, for as the stars revolve, a line joining the stars always passes through that center of mass (Figure 13-1). Thus, the true period can be determined by observations of the motion of the secondary relative to the primary.

The necessary observations are of the angular separation of the two stars (measured in seconds of arc) as well as of the orientation of the stars in the sky. At one time the two stars may be 5 seconds of arc apart with the secondary due south of the primary. At another time their separation may be 7 seconds of arc, and the secondary may be 10° east of south with respect to the primary.

The period can be determined quite easily from either the true

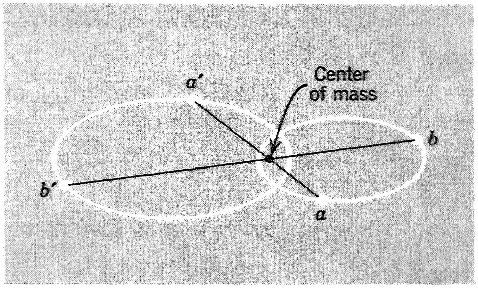


Figure 13-1 The motion of two stars in a binary system about their common center of mass.

orbit or the apparent orbit. The observed positions must be joined by the ellipse that best fits Kepler's second law, the law of equal areas. From this ellipse the period can be determined.

To determine the mean separation of the two stars in astronomical units, however, we must know how far the stars are from the sun. If they are close enough we can measure their heliocentric parallax and from it obtain the distance. More often than not, however, the stars are too distant for parallactic measurements. When they are, we must resort to a study of their spectra and determine their distance by spectroscopic parallax. When we know the distance and the apparent size of the true orbit in seconds of arc we can obtain the actual size of the orbit and thus the mean separation of the two stars in astronomical units. The period and the mean separation enable us to find the sum of the masses.

But we have only the sum of the masses. In order to obtain the ratio of the mean radii of each orbit we must follow the motion of *each* star as it moves relative to the background stars; that is, we must observe the proper motion of each star. The best example is the proper motion of Sirius, an A0 star, and its white-dwarf companion. In fact it was by following the proper motion of Sirius that its faint companion was discovered, for the motion of Sirius is not a straight but a wavy line. Figure 13-2 shows not only the proper motion of each star but also gives the relative orbit of the secondary about the primary as well as the orbit of each about their common center of mass. The inclination of Sirius is not 0° as in our earlier ideal example; it is 43°.

The size and shape of each apparent orbit is enough, however,

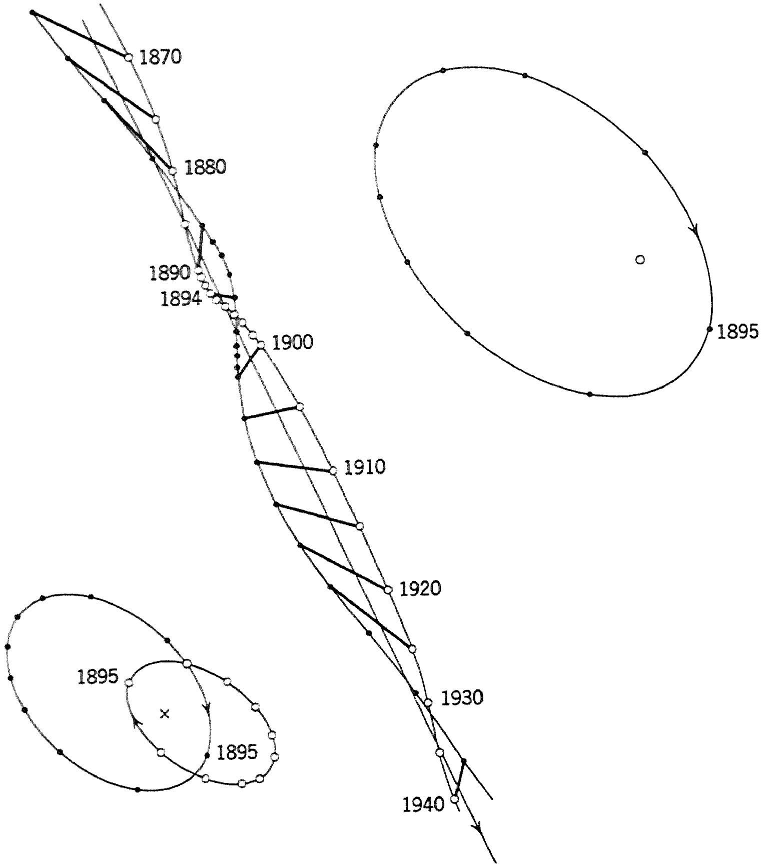


Figure 13-2 The motions of Sirius and its white-dwarf companion. Sirius is shown by a circle and its companion by a dot. Upper right: The apparent relative orbit of the companion about Sirius. Dots are at five-year intervals.· Lower: apparent orbits of Sirius and its companion about their common center of mass. Note that the center of mass (the focus of the elliptical orbits) is not at the focus of the apparent orbits. Center: the motions of Sirius and its companion relative to the background stars. The center of mass moves on a straight line. The scale is not the same for the three drawings. (Cecilia Payne-Graposchkin, *Introduction to Astronomy*, Copyright 1954, Prentice-Hall, by permission)

to determine the ratio of the mean radii, for this ratio is not altered by foreshortening. After we find the combined mass from Kepler's third law, the ratio of the radii shows the individual masses of Sirius to be 2.44 solar masses for the primary and 0.96 for the white-dwarf secondary. The true shape of an apparent orbit can be determined by the geometry of the situation since the inclination of the orbit foreshortens the distances involved in the same ratio. The true orbit is needed to get the true separation of the stars in astronomical units.

Thus for a visual binary, if the true relative orbit of the secondary about the primary can be found along with the binary's distance from the sun, the sum of the masses can be determined. The individual masses can be found only if the ratio of the mean orbital radii is known, and this ratio can be determined only if the motion of each star is referred to the background stars.

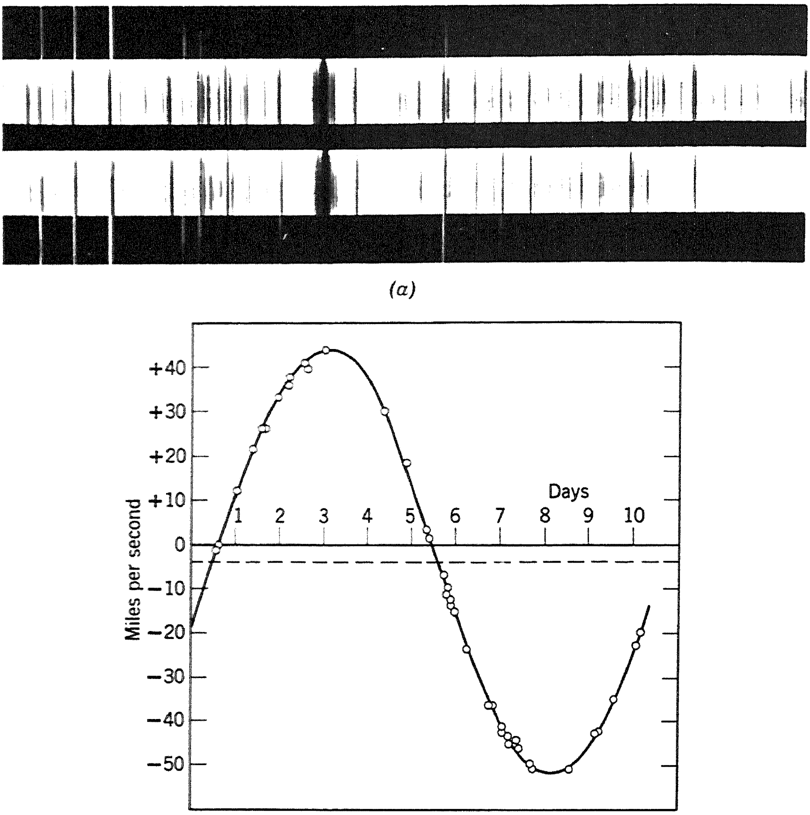
Spectroscopic Binaries

If the inclination of an orbit is large, the motion of the stars will be revealed by a periodic Doppler shift of their spectral lines. Since a line joining the stars always passes through the center of mass of the system, one star will be approaching while the other is receding from us, and their spectral lines will be shifted in opposite directions. Consequently, since the Doppler shift of the spectral lines is independent of the apparent separation of the stars in the sky, a binary system will always be disclosed by the periodic Doppler shift in the spectra of two stars, although these may be so close together that they cannot be resolved even with the largest telescope. A system that can be recognized by its spectrum is called a *spectroscopic binary*.

In Figure 13-3*a* we see two spectrograms of a binary system, each taken at a different time. In the top spectrogram the spectral lines are all shifted toward the violet, indicating that the brighter component of the pair is approaching the Earth; the lower spectrogram reveals that it is receding. The brighter component is so much brighter than the fainter component that the spectral lines of the fainter do not even appear. When the radial velocities are plotted on a graph against time, the velocity curve is obtained (Figure 13-3*b*).

The corresponding motions of the two stars are shown in Figure 13-4: (*a*) shows the brighter component receding from the Earth while the fainter is approaching; (*b*) shows the two stars traveling at right angles to the line of sight (at this time neither star shows a Doppler shift unless the entire system is moving relative to the sun); (*c*) shows the brighter component approaching while the fainter is receding.

If the two stars are of nearly equal magnitude, the spectral lines of the stars will be of nearly equal intensity, and both spectra will be visible. When the lines of one star are shifted to the violet the


(a)

(b)

Figure 13-3 (*a*) Two spectrograms of Iota Pegasi, a spectroscopic binary. The spectral lines of only the primary appear. In the upper spectrogram the lines are shifted to the violet, in the lower they are shifted to the red (Lick Observatory). (*b*) The velocity curve of Iota Pegasi. The dotted line indicates the radial velocity of the system. (Lick Observatory)

lines of the other star will be shifted to the red, but they will both be at their mean position at the same time. Therefore, as the stars continue to revolve about their common center of mass, the spectral lines of each star will oscillate back and forth across their mean position so that we sometimes see double lines and sometimes only single lines.

Since the velocities of the two stars are proportional to their distances from the common center of mass and since these distances are inversely proportional to the individual stellar masses, the ratio

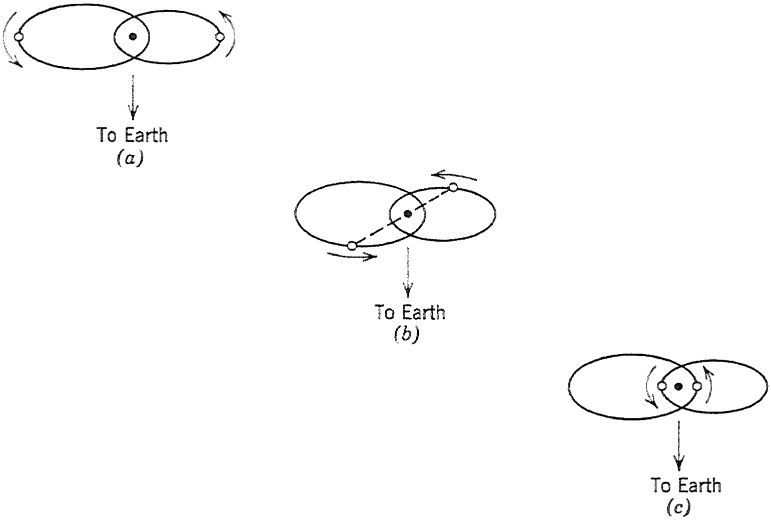


Figure 13-4 The motion of the two components in a spectroscopic binary. In (*a*) the primary star is receding and therefore its spectral lines are shifted to the red and those of the secondary (if visible) are shifted to the violet. The two stars in (*b*) are traveling in a direction perpendicular to the line of sight and thus their spectral lines are not shifted from their mean position. (*c*) shows the primary approaching and the secondary receding. The orbits are coplanar.

of their maximum velocities is inversely proportional to the ratio of their masses. We can therefore obtain the ratio of the masses when the spectral lines of both components are visible. But unless the binary system is actually a visual binary we cannot obtain their individual masses, for we cannot determine the inclination of the orbit. Hence we cannot determine the mean separation of the two stars, without which we are unable to determine the sum of the masses.

If one of the components of a spectroscopic binary system is much more luminous than the other, we see the spectral lines of only the more luminous star shifting back and forth. We can obtain its velocity curve but we can learn very little about the fainter component. We are nevertheless aware of the fainter star's presence, because the brighter component reveals a periodic shifting of its spectral lines without the same variation in brightness that occurs with the Cepheid. If the velocity curve of a Cepheid were mathematically treated as though it were a spectroscopic binary with only one velocity curve observable, the supposed fainter component would have to be inside

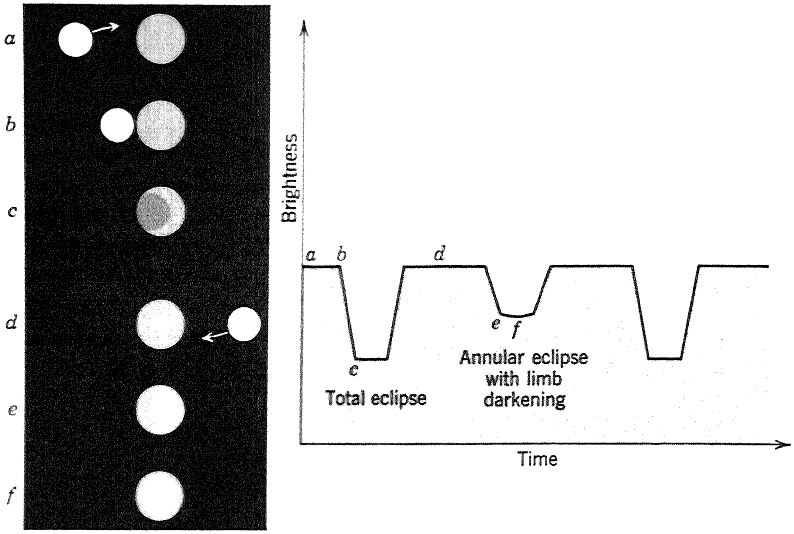


Figure 13-5 Eclipsing binary with total and annular eclipses showing the relative positions of the stars at various phases of the light curve.

the Cepheid. There are spectroscopic binaries, however, that have a Cepheid as one of their components.

Eclipsing Binaries

There are spectroscopic binaries with inclinations so large that one star passes in front of the other, causing an eclipse. These are called *eclipsing binaries.* As with the moon eclipsing the sun it is possible for the star nearest us to eclipse all or only a part of the star behind, and we can witness a total, a partial, or an annular eclipse. The type of eclipse can be determined from the light curve—the record of brightness of the system plotted against time.

Figure 13-5 shows a binary system with an inclination of 90°, in which the larger but fainter star can completely hide the smaller brighter star causing a total eclipse. An annular eclipse occurs when the smaller star passes between us and the larger star.

In order to descripe the shape of the light curve, let us assume the larger star to be stationary with the smaller one revolving about it. When the stars are not eclipsing we receive the maximum amount of light and the system is at its brightest (Figure 13-5*a*). The partial phase of the total eclipse begins when the smaller star just begins to go behind the larger star (*b*). The amount of light reaching the Earth starts to decrease until the smaller star is completely hidden from our view (*c*) and the total eclipse begins. The bottom of the light curve for a total eclipse is flat, for during this time we receive light only from the larger star. The total eclipse ends when the smaller star first comes into view from behind the larger star, and the amount of light reaching the Earth then begins to increase until both stars are completely in view (*d*). Actually all eclipsing binaries are so close together that we cannot resolve them into separate stars; thus the words "in view" refer to the stars as they are drawn in Figure 13-5.

After the total eclipse the smaller star revolves in its orbit to a point at which it goes in front of the larger star. When it begins to hide some of the larger star, the partial phase of the annular eclipse begins. The amount of light reaching the Earth decreases until the smaller star conceals an area of the larger star equal to itself (*e*). Since the area of the larger star that is concealed during an annular eclipse is always the same, namely, the area of the smaller star, the amount of light received during that phase is constant provided that the eclipsed star does not exhibit limb darkening (see p. 225f).

In an annular eclipse the amount of light that the smaller star blocks depends not only on its area but also on the temperature of the larger star. If the larger star does exhibit limb darkening, its limb is cooler and radiates less light per unit area than does the center of the disk. When the full annular eclipse is just beginning the smaller star conceals a region near the limb of the larger star (*e*). As the smaller star moves directly in front of the larger star (*f*) it continues to cover an equal area, but this area radiates more light than the limb. Therefore, if the larger star exhibits limb darkening, the bottom of the light curve of an annular eclipse is not flat but dips slightly in the middle when the smaller star covers up the more luminous central portion of the larger one.

In most eclipsing binaries the larger star has the lower surface temperature and radiates less light per unit area than the smaller star. Since the area eclipsed is the same (namely, the cross-sectional area of the smaller star) no matter which star is in front, more light is cut from our view when the smaller, hotter star goes behind, and the light curve may have two unequal minima. The deepest minimum, which occurs when the smaller, hotter star is eclipsed, is called the *primary minimum.* The *secondary minimum* occurs when the smaller

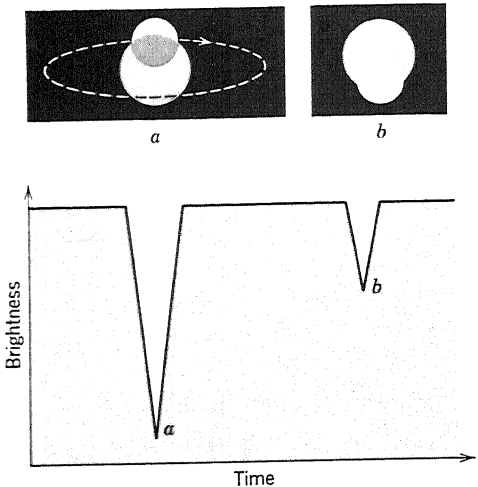


Figure 13-6 The light curve for a partially eclipsing binary.

star eclipses the larger, cooler one. If the two stars have the same surface temperature the two minima will have the same depth.

A partial eclipse results when the smaller star is never completely hidden by the larger (Figure 13-6). Thus the area of the smaller star concealed by the larger one always changes and the bottom of the minimum of the light curve is not flat. The light curve reaches a minimum when the maximum area of the eclipsed star is concealed, but at this instant the eclipsed star begins to move out from behind the eclipsing star, and the amount of light we receive increases once more. A partially eclipsing binary system is of less value to the astronomer than a totally eclipsing one, for the amount of area concealed in a partial eclipse is uncertain.

Total eclipsing binaries in which both stars are of about equal luminosity are extremely valuable to the astronomer, for they yield a wealth of information. If the stars have nearly the same luminosity, the system is also a spectroscopic binary with the spectra of both stars visible. Thus we can obtain the velocity curve for each star moving in its orbit. Since the inclination can be determined (for totally eclipsing binaries it is very nearly 90°), the maximum velocity of each star yields the actual velocity (in miles per second) of the star in its orbit. Most of the orbits of eclipsing binaries are nearly circular and hence the velocity of the star in its orbit is nearly constant. The actual velocity multiplied by the period (the time for one complete revolution) thus gives the circumference of the orbit. Once the circumference of each orbit is known, it is easy to find the orbital radii. With this information it is possible to determine the masses according to the method used with visual binaries. The period is known from the light or the velocity curve, the mean separation is the sum of the mean orbital radii, and since the radius of each orbit is known the ratio of the radii is easily determined.

For close binaries the duration of the eclipse depends on two factors: the size of the larger star and the velocity of each star in its orbit. The velocities can be determined if both spectra are visible, and by means of these velocities we can find the diameter of each star. The diameter of the smaller star is equal to the time required for it to cross the limb of the larger star (the time from the beginning to the end of the partial phase of a total or annular eclipse) multiplied by the velocity of the smaller star relative to the larger. The diameter of the larger star is equal to the time it takes the smaller star to traverse one diameter of the larger star (from the beginning of the partial eclipse to the end of the total eclipse) multiplied by the velocity of the smaller relative to the larger.

It is not uncommon to find one of the components of an eclipsing binary to be a red giant. Red giants, it will be recalled, are inflated stars with extended atmospheres. As the smaller hotter star goes behind the red giant, its light must travel through the atmosphere of the cooler companion, yielding of course a changing spectrum that gives us information about the atmosphere of the giant.

More than half of the 3,000 or so known eclipsing binaries have periods of less than 10 days, and more than three-fourths of them have periods of less than 100 days. Those stars with the shortest period are usually early type stars, mostly O- and B-types. A short period means that the stars are close together. Some of the eclipsing binaries are so close together they seem almost to touch.

Dwarf Novae

There is increasing observational evidence to support the theory that dwarf novae—and perhaps novae as well—are actually close binary stars. More than 20 dwarf novae have been shown to be spectroscopic binaries, several of which are also eclipsing binaries. The members of these close binary stars are interesting and their interaction must somehow lead to the recurrent outbursts of energy.

The best explanation of the dwarf nova hinges on the interaction of the gravitational fields of the two stars. The gravitational field

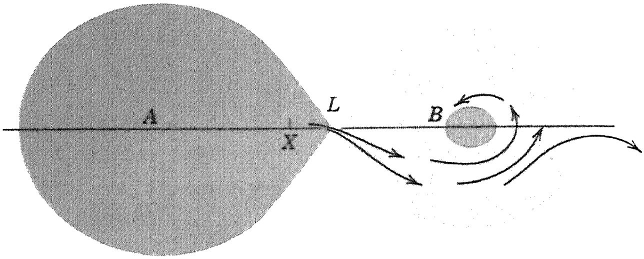


Figure 13-7 A proposed model of a close binary star resulting in a dwarf nova.

of a single star is, according to Newton's law of gravity, symmetrical in every direction. But the gravitational field which results from the proximity of the two stars of a close binary system is not symmetrical in every direction. There is a point between the two stars where the net gravitational force is zero; an object placed there with zero velocity would not fall in any direction. This point, called the *Lagrangian point L*, is actually the intersection of two closed surfaces the cross-sections of which are shown in Figure 13-7. These surfaces are called zero-velocity surfaces. A particle with zero velocity in lobe *A* cannot move into lobe *B*, and vice versa. But a particle with zero velocity at the intersection *L* of these two lobes may go either direction. If given a push towards *A* it will move into that lobe; a push towards *B* will move it in that direction.

Examination of the spectra of the two stars of dwarf novae indicates that their evolution is different from that of a single star. One of the stars in each observed dwarf nova is a small hot blue star; the other is a large cool red star. The large red star just fills its lobe defined by the zero-velocity surface. Gas appearing at the point *L* spills over onto the small hot blue star. But both stars are of nearly the same mass, and presumably the same age, so how can they be so different in their structure? How can one be an expanded cool red star and the other a small hot blue star?

A solution to this puzzle has been proposed by R. P. Kraft of the Mount Wilson and Palomar Observatories. Kraft suggests that originally the binary was composed of two stars of unequal mass. The more massive star burned its hydrogen in the core faster and so began to expand sooner. By expanding, it filled the lobe on its side, but gases—mostly hydrogen—could and did leak through the Lagrangian point onto the less massive star. By losing hydrogen gas, the more massive star evolved more rapidly, since evolutionary processes depend to some extent on the percentage of hydrogen. By acquiring hydrogen gas, the less massive star evolved more slowly. The originally more massive star evolved into a star approaching a white dwarf; the originally less massive star expanded and filled its lobe enclosed by the zero-velocity surface. This is the stage in which we now find the stars.

Gases now cross the point *L* from the originally less massive star and pour onto the star which is essentially a white dwarf. Presumably, the gases pouring onto the white dwarf cause it to erupt occasionally. If this is true, then the frequency of eruption must depend to some extent on the rate at which the gases pass through point *L*. There are many details yet to be filled in before this theory is able to explain all of the observed features of dwarf novae; in fact, the precise mechanism of eruption is not yet clear. The theory, however, has provided a direction for new observational research; more data are being collected, and a better theory will emerge. Out of all of this will come a better understanding of stars and their complex properties.

Beta Lyrae

Many close binary stars exchange gases. One very famous binary, Beta Lyrae, is almost a classical example of a close binary star with streaming gases. These gaseous streams can be detected in several ways. A few spectral lines, usually of hydrogen or helium, or both, show an increased intensity during certain phases of the period, indicating an uneven distribution of these gases about the system. Or these lines may show a Doppler shift that does not conform to the velocity curve of the system, indicating that these gases are moving with respect to the stars. Streaming gases may also be revealed by irregularities in the light curve.

Spectrograms of Beta Lyrae have been studied thoroughly by Struve, Kuiper, and other astronomers. Struve has chosen some of his spectra and arranged them as in Figure 13-8. The spectrograms have been aligned by the fairly faint but sharp line of Ca II (calcium ionized once) whose wavelength is 3934 Å. This line is not caused by the star but by gases between the star and the Earth; thus its Doppler shift depends on these interstellar gases and not on the motion of the star. By arranging the spectrograms according to this line, we see that the other lines (those caused by the star) change their positions throughout the period. This is the Doppler shift that yields the velocity curve.

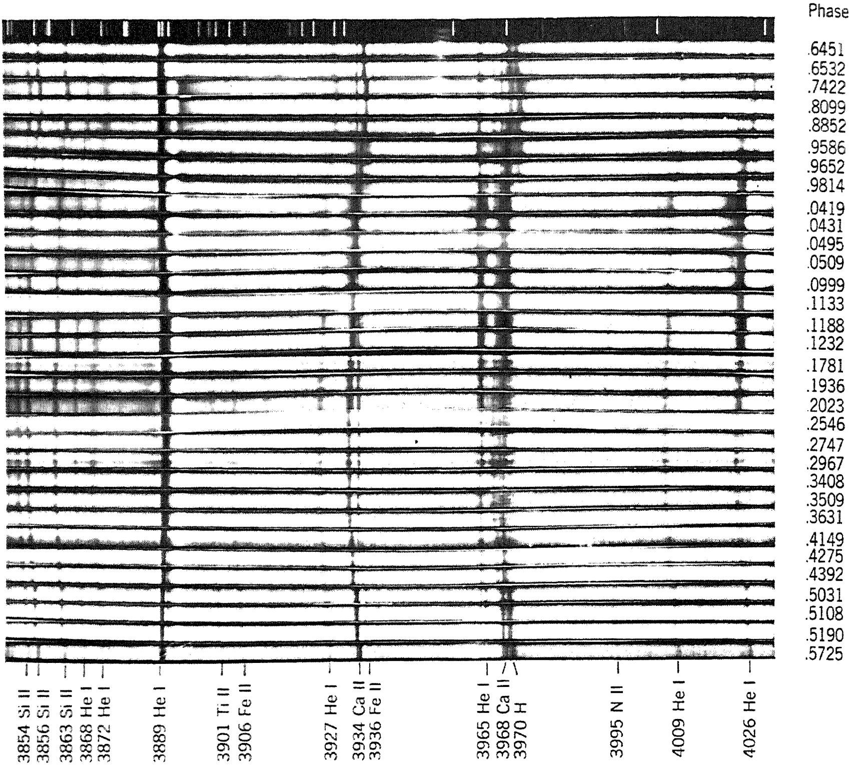


Figure 13-8 A sequence of spectra of Beta Lyrae that have been aligned by the interstellar line of Ca II 3934 Å. The wavy motions of the stellar lines result from the varying Doppler shift of the spectral lines as the primary component revolves in its orbit about the center of mass of the system. Notice that the emission component of the helium line, 3889 Å, has a nearly constant radial velocity. (Courtesy of O. Struve)

The line of He I (neutral helium), with a wavelength of 3889 Å, not only shows this Doppler shift but is seen in two parts, one in absorption and the other in emission. It can also be seen that the intensity of the helium line (and others as well) changes. The apparent increase in intensity near phase 0.04, however, is caused by an increase in the exposure time of the spectra necessitated by the star's going into minimum at this phase.

There are many peculiarities in the spectrum of Beta Lyrae that are not yet thoroughly understood, but continued studies reveal many interesting aspects of this binary system. The brighter component is a B8 star with a surface temperature of about 15,000°K, an orbital velocity of about 110 miles per second, and a period of revolution of 12.9 days. Thus the circumference of its orbit is the velocity multiplied by the period in seconds (1.1×10^6 seconds) or 1.2×10^7 miles. The radius of the orbit equals the circumference divided by 2π or about 2×10^7 miles. From a study of the duration of the eclipse the radius of the B8 star has been estimated to be about 3×10^7 miles. Since the radius of the orbit is smaller than the radius of the star itself, the center of mass of the system must be inside the B8 star. Its radius of 3×10^7 miles classifies the B8 star as a supergiant. The sun's radius is only 4.3×10^5 miles.

The spectral lines of the fainter component have never been seen, so it must be considerably less luminous than the B8 primary. From studies of the light curve it appears to have the luminosity of an F-type star and a surface temperature of 8,000°K. Since only the lines of the B8 star are seen in the spectra, a velocity curve for the fainter companion cannot be made, and there is considerable doubt about the masses of the two stars. It does seem, however, that at least one of the stars is at variance with the mass-luminosity relation (see p. 264).

Near the time of primary minimum the lines of hydrogen and helium are flanked by "satellite lines." The fact that these lines are fainter than the principal lines of these elements and have a Doppler shift different from that of the principal lines, indicates that they are not formed by the gases on the surface of either star. Just before primary eclipse the satellite lines appear on the red side of the principal lines; just after eclipse they appear on the violet side. The red satellite lines yield a radial velocity of 90 to 120 miles per second, and the violet satellite lines yield radial velocities of 50 to 220 miles per second.

The interpretation placed on these curious lines is that they result from streams of gases that flow from each star (Figure 13-9*a*). It can be seen that such streams of gas will yield their largest Doppler shift just before and after the F star eclipses the B8 star.

There are also hydrogen and helium lines that seem to originate from a shell of gas that does not participate in the motion of the B8 star. Instead, they show a rather constant velocity of approach of roughly 60 miles per second, indicating an expansion of gases around the entire binary system. Such gases, according to Kuiper, may take the shape of a spiralling watch spring (Figure 13-9*b*).

Most of the binary systems in which gases stream from one star

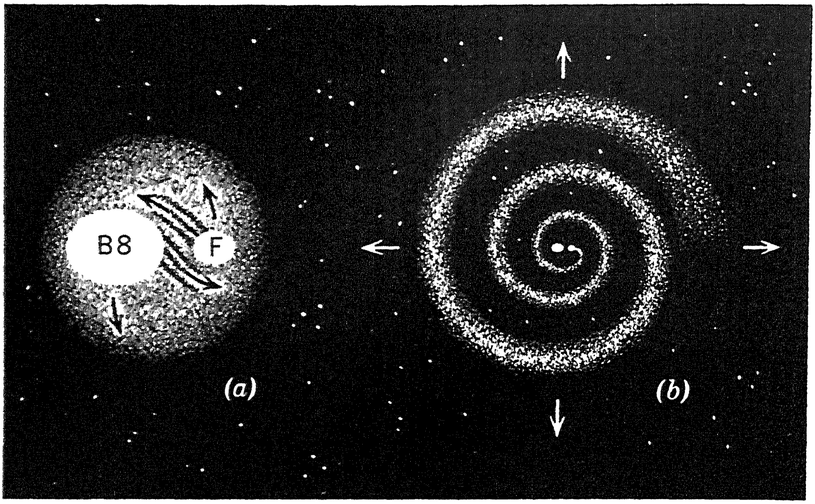


Figure 13-9 The most successful interpretation for the unusual spectral features for Beta Lyrae assume that (*a*) gases stream from both stars, enveloping them in an extended atmosphere of sorts, and (*b*) that gases are ejected from the system which, because of the rotation of the system, spiral out into space. (After O. Struve)

to the other are close binaries of early spectral type. The loss of mass that results from the gaseous streams should be reflected in a change in the period of revolution of the system. And indeed the period of Beta Lyrae is increasing at the rate of 19 seconds each year.

Stars or Planets?

It seems significant that the average distance of spectroscopic binaries is about 20 A.U. Furthermore, it seems reasonable to presume that whatever forms in the nebular disk about a star depends upon the conditions of that disk. It is, consequently, reasonable to assume that either stars or planets may form in such a disk. Since double stars have formed, it is again reasonable to assume that planets have also formed about other stars.

There is, in fact, convincing evidence that planets are revolving about other stars. Work towards discovering such planets has been carried out at Sproul Observatory, Swarthmore College, under the directorship of Peter van de Kamp. There are at least three stars whose position is close enough to the sun so that studies of their proper motions have revealed an unseen companion: 61 Cygm, Lalande 21185, and Bernard's star. The companion of Bernard's star revolves in a period of 24 years, travels in an orbit with a mean radius of 4.4 A.U., and has a mass 0.0015 times that of the sun. Its mass is only 1.5 times that of Jupiter (Figure 10-3).

A gaseous sphere with a mass less than 0.07 solar mass cannot generate a high enough temperature in its core to sustain thermonuclear reactions. Therefore, a body with a mass less than 0.07 solar mass ought to be classified as a planet! The companion of 61 Cygni has a mass of 0.01 solar mass, so does the companion of 21185 Lalande. If each of these stars has at least one planet revolving about it, it most likely has several. What about more distant stars whose proper motions are so slight as to prohibit fruitful studies?

From a study of the luminosity function (see p. 266), Harrison Brown, of the California Institute of Technology, has drawn conclusions about the possibilities of planets accompanying other stars. The luminosity function clearly indicates that the number of stars in the solar neighborhood increases as stars of decreasing luminosity are considered, until an absolute magnitude of +13 is reached. As still fainter stars are considered, the number of stars decreases, not increases. The reason for this decrease in the number of stars could be that these stars are simply too faint to be seen from the Earth; or it could be that a star with an absolute magnitude of +13 has a mass approaching 0.07 solar mass. Perhaps the luminosity function falls off because at about absolute magnitude +13 the objects which form may or may not be stars—the luminosity function may have reached the stage when planets should be counted.

By assuming that planets do indeed form from objects of mass less than 0.07 solar mass, Brown has extended the luminosity function to speculate about how many planets there are in the galaxy. His conclusions include the statement that "virtually every main-sequence star should have a planetary system associated with it." Furthermore each star should have, on the average, two terrestrial planets at such a distance that water remains in the liquid phase. Any one of these planets could support life. Our planet does!

These examples illustrate the value of binary stars to the astronomer. All our knowledge of stellar masses, which has proved so vital in the theories of stellar evolution, is based on the masses obtained from binary systems, for it is possible to measure the mass of a star only if it is a member of a binary system. Basic information concerning stellar sizes and densities has been learned from the eclipsing binaries.

The fact that limb darkening can be observed on stars in an eclipsing system even though we cannot see the disk of the star has extended our knowledge of stellar surfaces. Valuable information concerning stellar evolution has been obtained from binaries, including the recognition of stars that do not radiate visible light and therefore could not be detected if they were not members of a binary system. Much of the information on white dwarfs has been obtained from the white dwarfs that are members of binary systems. In view of all this it is not surprising that binary star are studied assiduously, especially eclipsing binaries with the spectra of both stars visible.

CLUSTERS

Binary stars may travel alone through space or they may be accompanied by other stars. Epsilon Lyrae is composed of two pairs of binaries. The naked eye can see two stars close together, but the telescope reveals that each of these is a visual binary, one star of which is a spectroscopic binary. Castor, the brightest star in the constellation of Gemini, is seen as one star to the unaided eye, in the telescope it appears as two, but with the spectrograph each is revealed to be a spectroscopic binary. There is another spectroscopic binary close to Castor that is moving with the same speed and in the same direction as Castor and therefore must be a physical member of the same system, bringing the total to six.

The gregarious nature of stars is revealed most startlingly, however, in the clusters, which contain as many as 100,000 stars all held together by the mutual gravitation of the stars in the cluster.

Clusters can be divided very nicely into two distinct kinds: *open clusters* (also called *galactic clusters*—a term easily confused with clusters of galaxies, so we prefer the more descriptive term) and *globular clusters.* The most obvious differences between these two types are the number of stars contained in each as well as their general shape and appearance, and location in the galaxy.

Open clusters have from around 20 to several thousand stars in a loose assemblage of rather indefinite shape. Globular clusters, on the other hand, may be composed of as many as 100,000 stars held together in a rather compact, spherical, or ellipsoidal shape.

Open Clusters

There are several open clusters that are easily visible to the naked eye. The most famous, perhaps, is the Pleiades or the Seven Sisters.

Many stars in the constellation of Taurus form another open cluster called the Hyades. The Big Dipper (in Ursa Major), is also an open cluster of some 40 stars (the two endmost stars are not members of this cluster). We are so close to the Ursa Major cluster that one of its members, Sirius, appears in quite a different part of the sky. The sun, however, is not a member of this cluster or of any other known cluster.

Without exception the open clusters are found very close to the central plane of our Milky Way galaxy. This striking conformity makes it difficult to identify open clusters at very great distances from the sun. They simply blend in with the profusion of background stars in the Milky Way.

Members of a cluster may be identified by their motion. If the cluster is close to the Earth, a common proper motion will identify the stars that physically belong to the group. If the proper motion of a particular star in the group is different, it may be either behind the cluster or between the cluster and the Earth. Although each star in a cluster moves with respect to the cluster, this velocity is small compared to the velocity at which the cluster is moving through space.

The Hyades cluster illustrates the principle of common proper motion better than any other because it is close enough to reveal good proper motions and yet compact enough to appear as a cluster on a photograph. From Figure 13-10 it can be seen that all the stars in the cluster are moving toward a common point called the *convergent point,* which indicates that the cluster is moving away from the sun. As the cluster recedes from us, its apparent size will decrease, and the paths of the individual stars converge like railroad tracks leading across flat country. The average radial velocity of the stars in the cluster is about +21 miles per second, which confirms our observation of the cluster's recession.

The brightest star in the constellation of Taurus is the binary, Aldebaran. Its proper motion, however, indicates that it is not a member of the Hyades cluster. Its lack of membership becomes even more convincing when we realize that its parallax gives a distance of only 58 light years, and the cluster is 130 light years away. All the stars that belong to the Hyades are included in a volume of space about 33 light years in diameter.

Proper motion studies of the more distant clusters are impractical as a means of identifying cluster members, however, for the greater

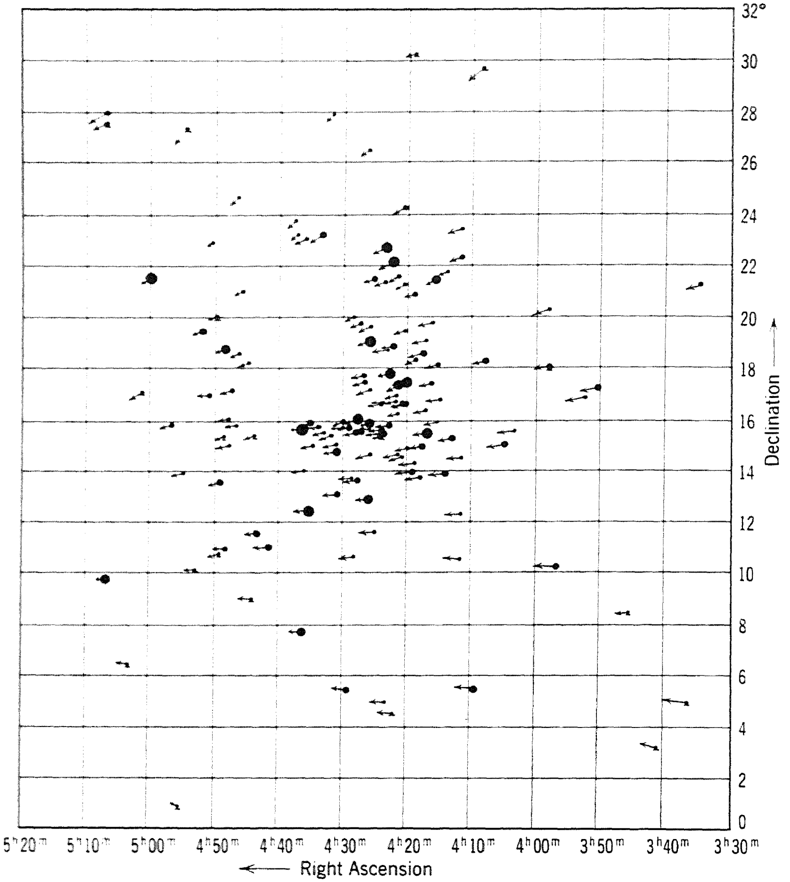


Figure 13-10 The proper motions of the stars in the Hyades cluster converge to a point, indicating that the cluster is receding from us. The size of the dots on this drawing by Van Bueren, of Leiden Observatory, indicates the magnitudes of the stars. The arrows show the proper-motion displacement expected in the course of the next 18,000 years. (Reprinted by permission of the publishers from Bart J. and Priscilla F. Bok, *The Milky Way*, Cambridge, Mass., Harvard University Press, Copyright 1957 by the President and Fellows of Harvard College)

distances mean that the proper motions are correspondingly reduced. It is nevertheless possible to identify members by radial velocity measurements. For example, the radial velocity of Aldebaran is +30 miles per second and not +21, which is the radial velocity of the Hyades cluster.

With the utilization of the photoelectric cell for obtaining very accurate measurements of the brightness and color of a star, it has become convenient and instructive tc make an H-R diagram of stars in a cluster. The stars in the field that fit the expected pattern are considered members of the cluster in good standing; the stars that appear on the diagram in a position quite different from the bulk of the cluster stars are rejected from membership. Other criteria, such as radial velocity and spectroscopic parallax, may be employed to differentiate those stars whose membership is in doubt.

The use of the photoelectric cell to determine color replaces the spectral sequence in the H-R diagram. Since both the spectral type and the color of the star are dependent on the temperature, they are often used interchangeably. However, since spectral classification depends also on the pressure of the gases in the atmosphere of a star, a giant and a dwarf of the same color belong to slightly different spectral classes.

The color of a star is rated by its *color index,* a number indicating the relative brightness of the blue and the yellow regions of the spectrum.* By use of appropriate filters in front of a photocell, measurements of apparent magnitude can be made which will approximate measurements made with the eye. These magnitudes are called visual magnitudes V. Other filters approximate magnitudes established by photographic means; these are called *photographic magnitudes* B, for blue. The visual magnitudes depend upon the color sensitivity of the eye which reaches a maximum in the green-yellow region of the spectrum. Photographic magnitudes are based upon the sensitivities of commonly used photographic emulsions which are most sensitive to the blue, violet, and the near-ultraviolet. A blue star will appear brighter in the photographic magnitude system than in the visual magnitude system, and the reverse is true for a red star. By agreement, the two systems have been adjusted so that the magnitude of an A0 star is the same in both.

The color index of a star is the difference between the photographic and the visual magnitudes. It is calculated by subtracting the visual magnitude V from the photographic magnitude B: $B - V$. Consequently, an O- or B-type star will have a negative color index; since it will appear brighter on the photographic system, its photographic magnitude will be numerically *less* than the visual magnitude. Any star later than A0 (for example, a G-type star) will have a positive color index, for it will appear fainter in the blue and thus have a larger number on the photographic magnitude scale than on the visual scale. Table 13-1 gives the color indices of main-sequence stars of different spectral types. This table is based on an overall temperature of 5,730°K for the surface of the sun, a G2 star. This temperature includes limb darkening to keep it in line with stellar temperatures. The color index of the sun is +0.65.

---

* There are other color indices relating other regions of the spectrum; for example, the ultraviolet to the blue.

Spectral Type	*Temperature*	*Color Index*
B0	25,000°K	−0.32
B5	15,600	−0.16
A0	11,000	0.00
A5	8,700	+0.15
F0	7,600	+0.30
F5	6,600	+0.44
G0	6,000	+0.60
G5	5,520	+0.68
K0	5,120	+0.82
K5	4,400	+1.18
M0	3,600	+1.45

Since the color index can be determined so much more quickly and more objectively than the spectral type, it is common practice to plot a color-luminosity diagram for a star cluster. A color-luminosity diagram represents essentially the same thing as the spectral-luminosity (H-R) diagram.

Not only has the color-luminosity diagram enabled astronomers to identify the members of distant clusters, it has also yielded valuable information concerning the evolution of stars within the many clusters. Figure 13-11 shows a color-luminosity diagram for a number of open clusters. All the clusters have stars on the main sequence, the main branch in the diagram, but each cluster is represented by a different branch off the main sequence. The double cluster in Perseus (H and Chi Persei) has many very hot blue luminous stars. The hottest stars in the Pleiades, although very luminous, are not O-type stars. The hottest stars in NGC 188 are not much hotter than the sun. These differences are suspected to result from differences in the clusters' ages. The youngest, NGC 2362, is perhaps less than one million years old. In NGC 188, however, even the moderately

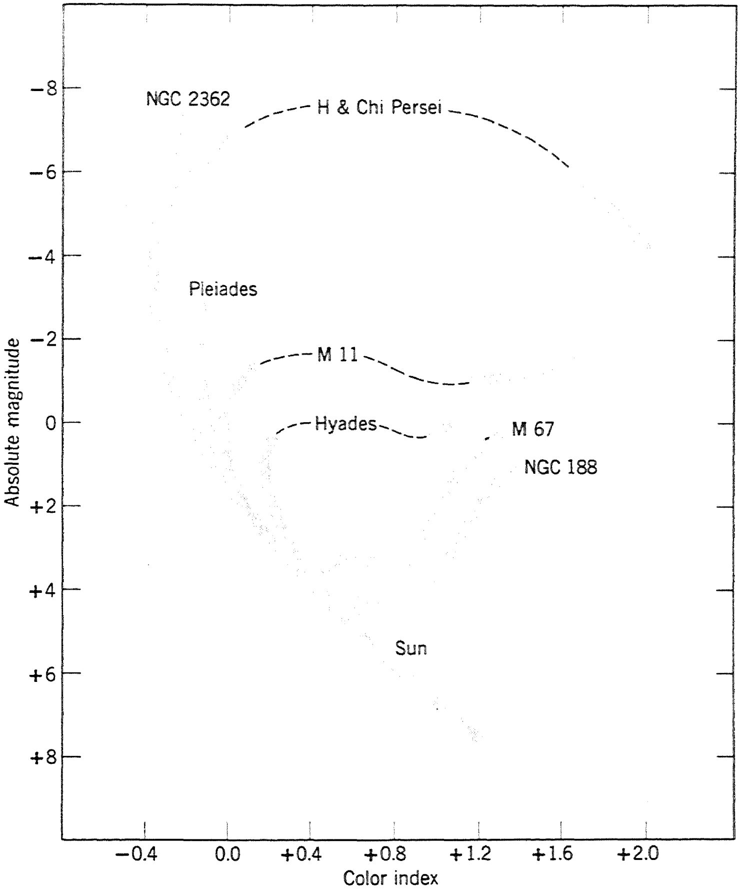


Figure 13-11 A composite color-luminosity diagram for several open clusters. (After Sandage)

massive stars are starting to evolve off the main sequence toward the red-giant stage; it must, therefore, be the oldest cluster on this diagram.

The Pleiades cluster (Figure 13-12) is 400 light years from the sun and, with its early B-type stars, is young on the astronomical scale. The time required for a B-type star of 9 solar mass to contract to the main sequence, deplete its core hydrogen, and start its expansion into a red giant is estimated to be about 21,000,000 years. This, then, should be the age of this open cluster. A color-luminosity diagram of the Pleiades alone, however, makes it clear that stars of only 0.2 solar mass have already passed through their initial contrac-

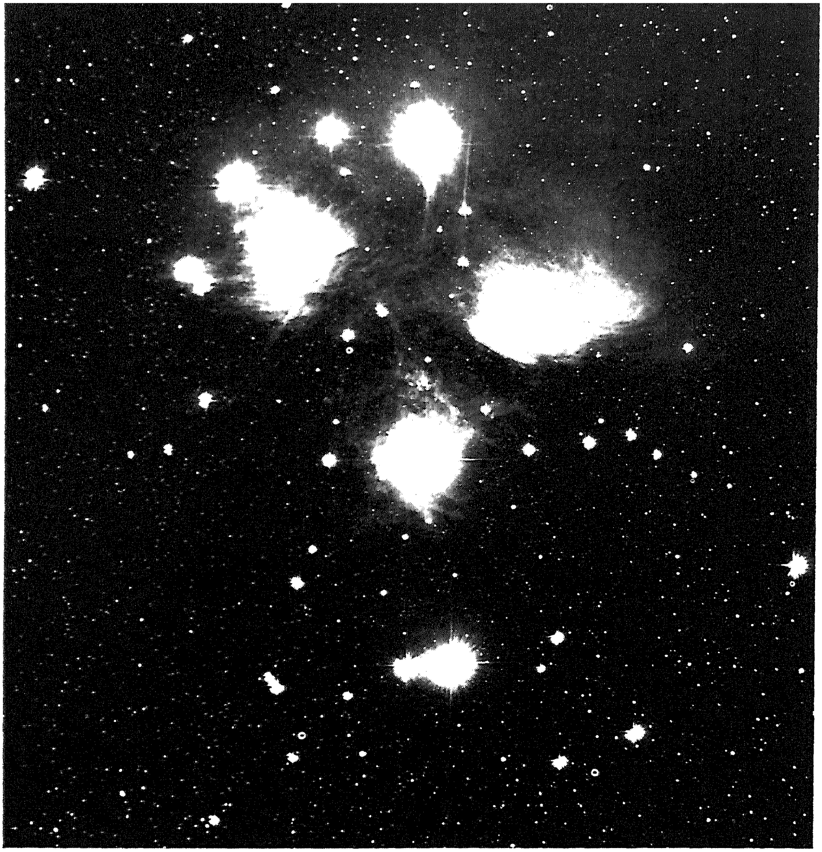


Figure 13-12 The Pleiades cluster is embedded in nebulous dust. (Lick Observatory)

tion stage and are essentially on the zero-age main sequence. Current theories of stellar evolution estimate that it takes roughly 600,000,000 years for a star of 0.2 solar mass to contract onto the zero-age main sequence. So how old is the Pleiades cluster, 21,000,000 years or 600,000,000?

Actually the discrepancy is not so great as it may appear. It has been suggested by Herbig that stars of small mass form from the nebular material before stars of large mass form. If the late-type main-sequence stars form first, they may be able to reach the zero-age main sequence in about the same time interval required for the early-type stars to contract onto the main sequence, burn their core hydrogen and start the expansion to the red-giant stage. Theory seems to be supported by observation. Additional evidence for the early formation of low-mass stars is seen in the Taurus-Auriga dark clouds which have a high population of faint red stars and no luminous blue stars.

All this indicates that the Pleiades, with roughly 300 stars, started to form 600,000,000 years ago and continued until the B-type stars finally formed. These very luminous stars, which radiate ultraviolet radiation copiously, have ionized the gases and literally blown all but the remaining vestiges of gas out into the galaxy. With this gas gone, stellar formation was brought to a halt. It is interesting to speculate that the Rosette nebula (Figure 12-3) with its cluster of very luminous stars at the center may be caught in the act of blowing the gases away from the central region. This may account for the ring-like structure of this very interesting aggregate of gas and stars.

All of these theories of star formation are well supported by the Orion complex. The Orion complex, it may be recalled, has some Herbig-Haro objects and an overabundance of T Tauri stars. The Orion region is also liberally populated by O- and B-type stars that form a very loose group called an *association*. An association is not a cluster but a large number of stars of similar type in the same region of the galaxy. The association of stars in the Orion region is embedded in a huge tenuous cloud of hydrogen gas some 1,500 light years from the sun. Only a small part of this cloud is visible to the naked eye as the Orion nebula (Figure 2-9), because only in its central region is there a group of very hot luminous stars. The entire cloud may have a diameter as large as 300 light years while the visible portion is about 20 light years in diameter. Radio observations indicate that the gas is distributed in a symmetrical pattern around the group of hot blue stars. The uneven appearance on optical photographs results from interstellar dust in the foreground.

Studies of the entire Orion complex suggest that it is expanding. The hot blue stars in the central portion (Figure 13-13*a*) cannot be very old, and their motion indicates that they all formed in an even smaller volume space some 10,000 years ago. Studies of the extent of ionization and the motions of the gas indicate the nebula itself has an age of some 10,000 to 20,000 years. Presumably it existed as a dark nebula long before that, perhaps a million years ago, during which time less massive stars formed, which now appear as the faint red stars in Figure 13-13*b*. However, once the very massive stars

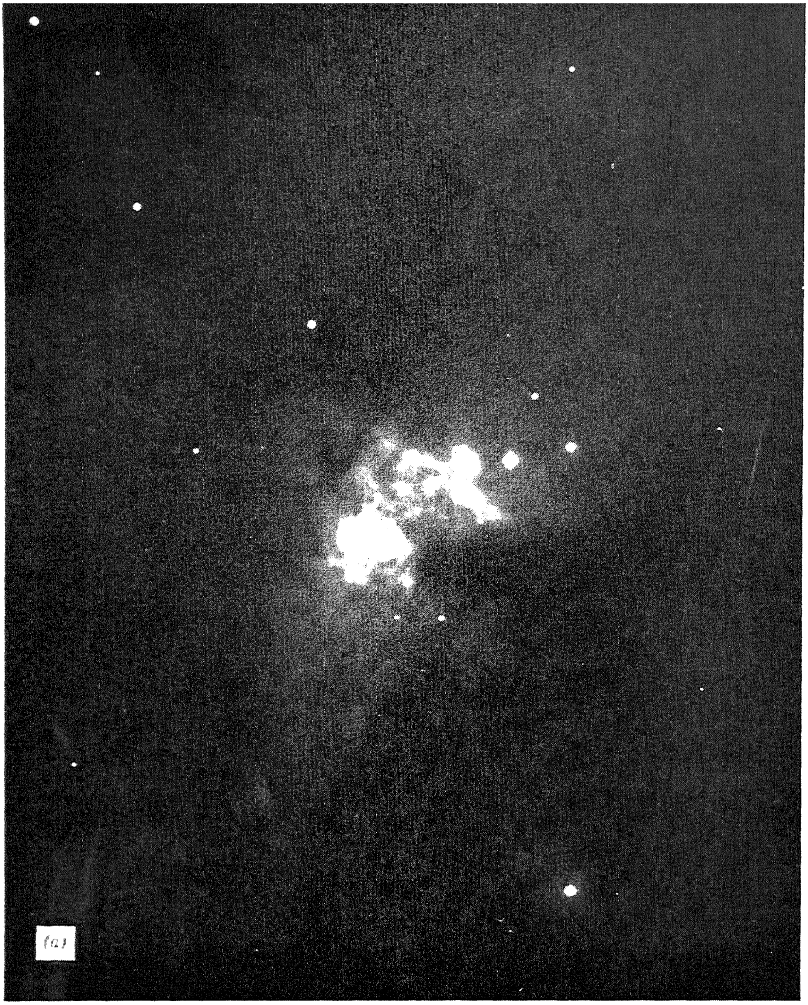
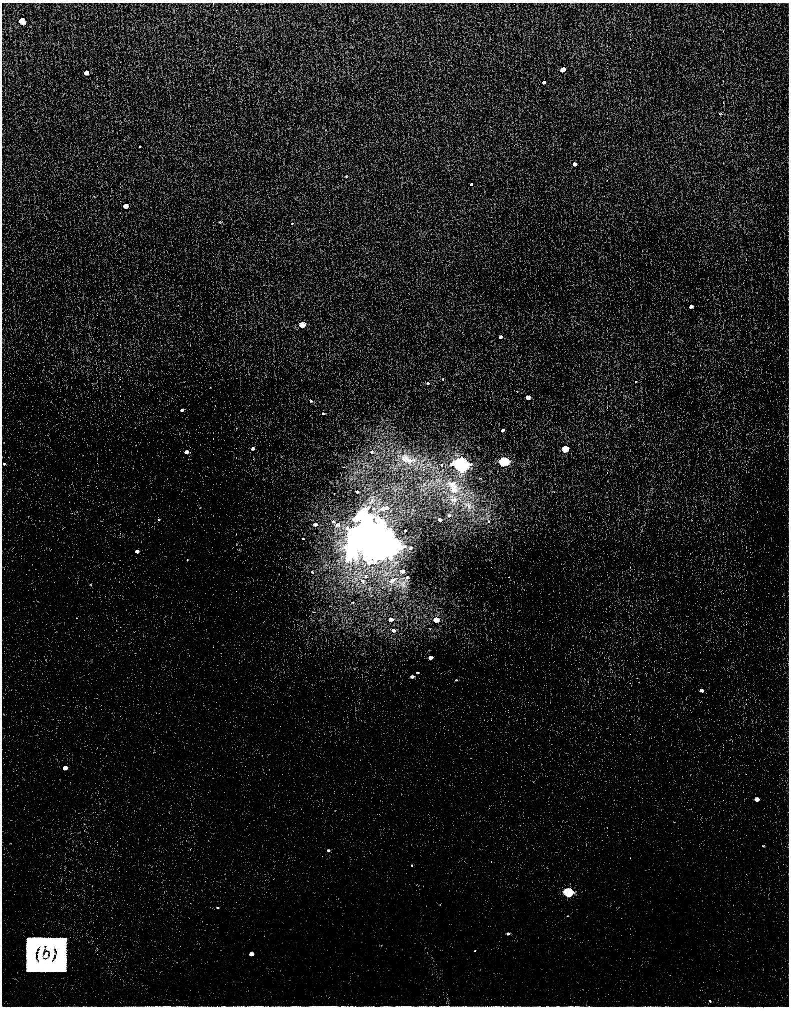


Figure 13-13 The Orion nebula (*a*) in ultraviolet light and (*b*) in infrared light. (Courtesy of G. Herbig of the Lick Observatory)

of the central part formed with surface temperatures ranging from 20,000° to 50,000°K, the gases became ionized and the expansion started. As the expansion continues, star formation in the central regions must decrease.

The color-luminosity diagram of the Orion association of stars shows that the most massive stars have not only had time to evolve onto the main sequence, but that they have burned enough of their hydro-



gen to start their evolutionary process which carries them to the right on the H-R diagram. The stars of moderate mass are on the zero-age main sequence, but the stars of very small mass have not yet reached the zero-age main sequence. Their slow contraction assures that they will reach the main sequence last of all—even if they started to form first. Their handicap does them no good!

The Orion complex, profusely populated with young stars all embedded in a matrix of hydrogen gas, convincingly supports the belief that stars are born of gaseous clouds.

It is not well understood why an association of stars forming in a nebula either breaks up shortly after formation or remains to become an open cluster. One suggestion deals with the percentage of the cloud that condenses into stars. The cloud, originally held together by the gravitational attraction of the gases that compose it, may condense largely into stars with only a little gas left over. In this situation most of the mass remains and the gravitational field remains strong enough to hold the stars in a more permanent cluster. If, on the other hand, a goodly amount of the cloud does not condense into stars, it is likely to get blown away by the particle pressure of those stars that have formed; consequently the total gravitational field will be reduced and the association is likely to break up sooner.

Globular Clusters

Unlike open clusters, globular clusters are characterized by extreme stability. Not only are they very compact, with many more stars than the open clusters, but their ages are much greater—in the neighborhood of 5 billion years or more. This follows from the observation that globular clusters contain many red giants and RR Lyrae stars. One globular cluster even contains a planetary nebula. Each of these objects indicates a much greater age than do the blue giants in open clusters. The relative ages of the two types of clusters are made clear by the differences between the typical color-luminosity diagram of each (Figure 13-11 for the open cluster and Figure 12-6 for the globular cluster).

Globular clusters contain few if any early-type stars on the main sequence. These stars, if they once existed in the globular clusters, have long since passed through their evolutionary track and may exist as white dwarfs too faint to be seen in the more distant globular clusters. The color-luminosity diagrams of globular clusters also differ considerably from those of the stars in the neighborhood of the sun. Most of the stars near the sun are main-sequence stars of types later than A0; few are red giants, RR Lyrae stars, or planetary nebulae.

Not only do globular clusters contain stars of a type different from those found in open clusters and in the neighborhood of the sun, but globular clusters are distinguished even further by their location in our galaxy.

There are no globular clusters close to the sun; the brightest, Omega

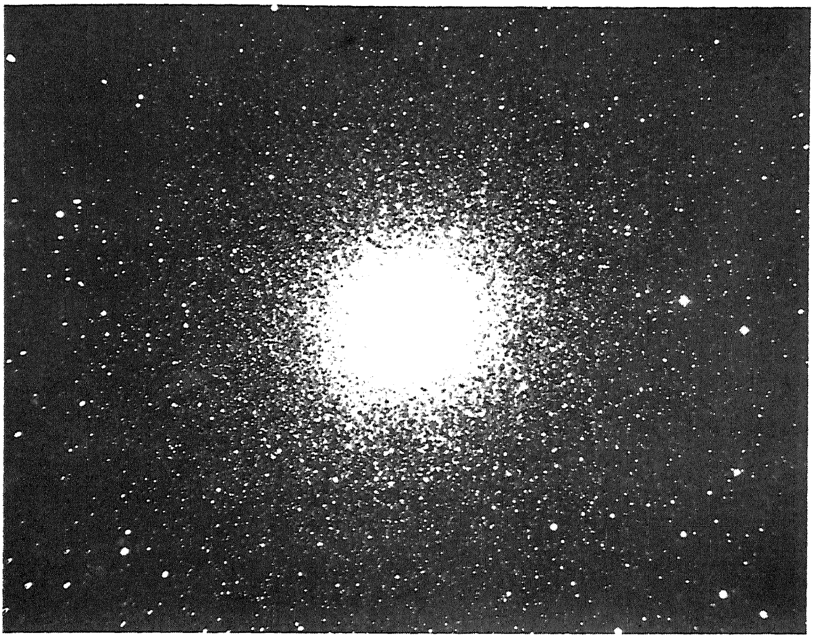


Figure 13-14 The globular cluster M13 in Hercules which must contain more than 50,000 stars. (Mount Wilson and Palomar Observatories)

Centauri, is 22,000 light years from the sun. Despite its great distance it is visible to the naked eye in the southern hemisphere as a faint patch of light in the sky, because it has many thousands of stars. The globular cluster in the constellation Hercules (Figure 13-14) is 34,000 light years away and can be seen with the aid of only a small telescope.

There are about 121 globular clusters known. Of these 30 (25% of the total) reside in a very small portion (2%) of the sky in the constellation Sagittarius and the rest are spread throughout the entire sky. This concentration in Sagittarius is especially significant, for the center of our galaxy lies in that direction. Since both the 30 globular clusters and the galactic center lie at about the same distance from the sun and in the same direction, these clusters must be concentrated about the center of our galaxy.

The remaining globular clusters lie at the fringes of our galaxy and form a nearly spherical halo concentric with the center and contrasting with the flatness of our galaxy (Figure 13-15). Open clusters, on the other hand, lie close to the plane of our galaxy.

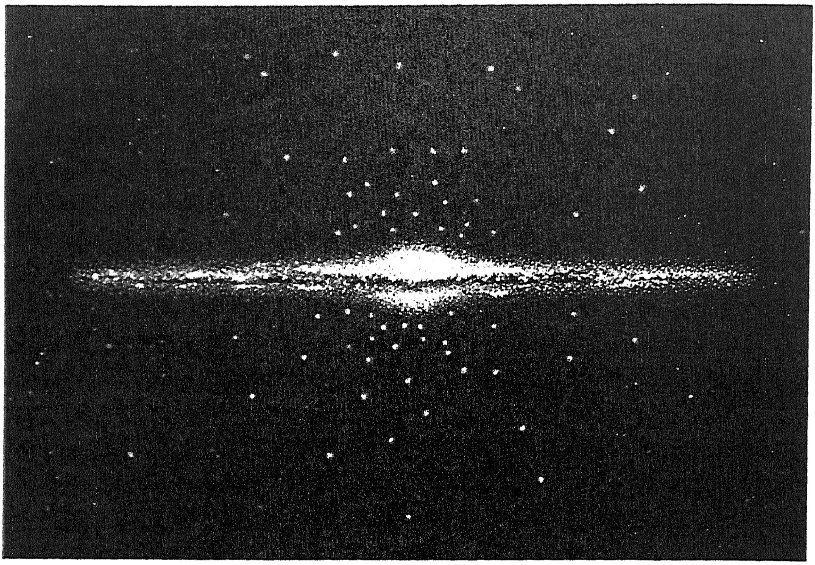


Figure 13-15 Our galaxy is surrounded by a halo of globular clusters.

Because of their location in our galaxy the clusters can be easily classified into the two main stellar populations. An open cluster generally consists of young, blue-giant stars, lies close to the galactic plane, and is therefore of Population I. A globular cluster consists of stars such as red giants and RR Lyrae stars that are well advanced in their evolution. In addition, it is either far from the galactic plane or near the galactic center. Consequently it is of Population II.

Since there are no globular clusters of Population I we can only assume that they were all formed billions of years ago at a time when our galaxy was quite different from what it is now. Similarly, since there seem to be very few extremely old open clusters we can only assume that open clusters are too unstable to enjoy a life of billions of years. But speculations on the age and evolution of our galaxy must wait until Chapter 15.

In apparent contradiction to theories of stellar evolution, the globular clusters, with their red giants and RR Lyrae stars, also have a few fairly hot blue stars on the main sequence. The color-luminosity diagram of M3 (Figure 12-6) reveals some of these early-type stars—those stars on the main sequence above the *turn-off point.* All the other fairly massive stars that once existed on the upper part of the main sequence have expanded and evolved to the right. Why should these few stars remain? After all, a star's location on the main sequence and the length of time it remains there are both dependent on its mass. Therefore, if these early-type stars were formed at the same time all the other stars in the cluster were formed, we see no reason why they, too, should not have evolved off the main sequence.

One possible solution to this dilemma is to assume that these stars, somehow, formed after the bulk of the stars in the cluster had formed. If so, they are younger and behind the rest in their evolution. In support of this, there is some evidence that globular clusters contain intracluster gases. The dark patch in the upper left of M13 (Figure 13-14) may be some of this gas, that presumably could have given rise to stars younger than the cluster. It is speculated that this intracluster gas could come from the more massive cluster stars that have already ejected gases to become white dwarfs.

It does not seem likely that intracluster gases could have been picked up from our galaxy, for globular clusters are never found associated with interstellar gases. Each is located in a region of the galaxy that is quite devoid of nebulae, whereas the more recently formed open clusters are very commonly associated with nebular gases. These gaseous clouds, along with interstellar gases, form a goodly portion of the mass of our galaxy. Although they conceal a large percentage of the stars, they have given us valuable information on the structure of our galaxy.

Basic Vocabulary for Subsequent Reading

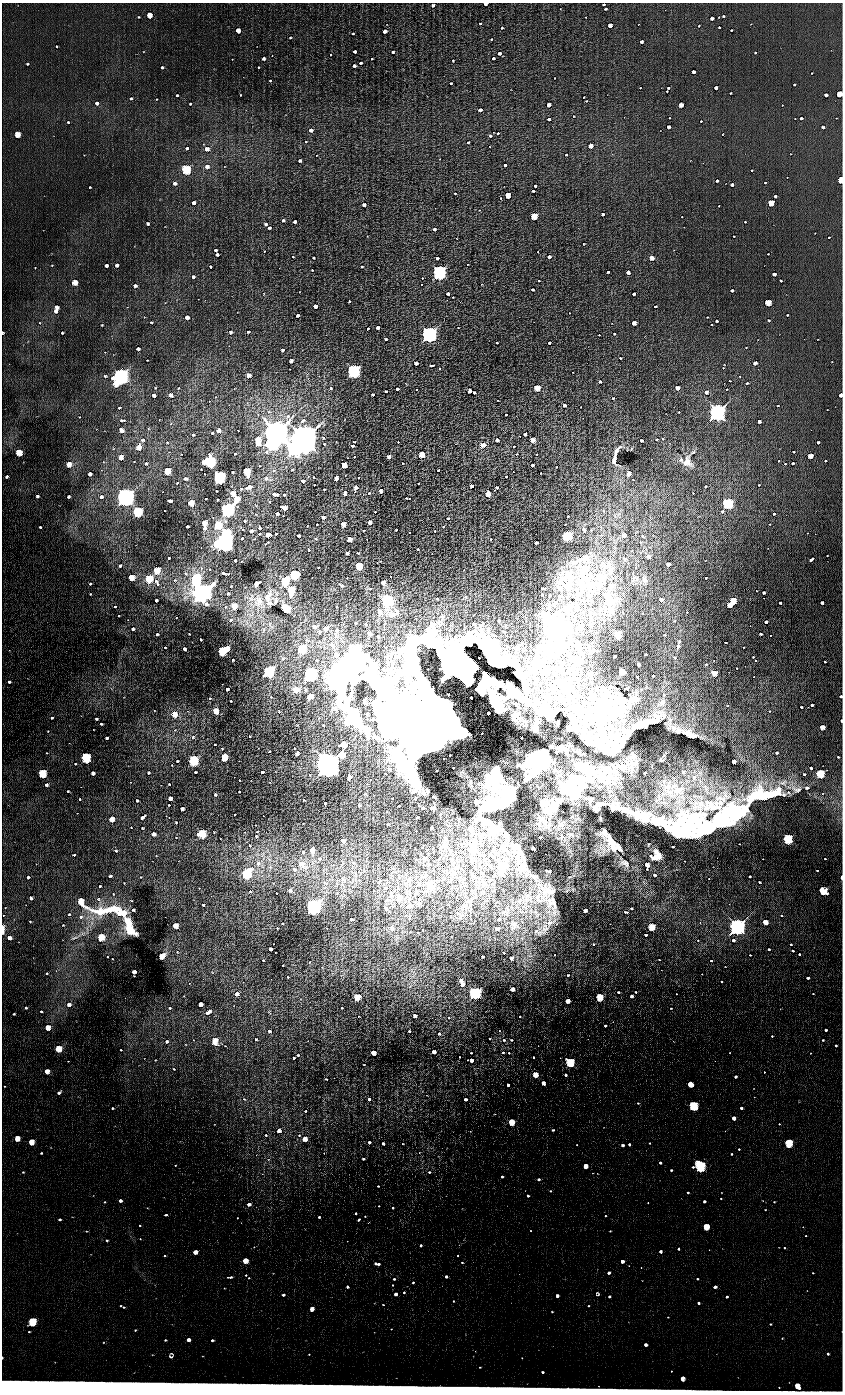
Apparent orbit
Color index
Eclipsing binary
Globular cluster
Open cluster
Photographic magnitude
Spectroscopic binary
Stellar association
True orbit
Visual binary
Visual magnitude

For Further Reading

Struve, Otto, and V. Zebergs, *Astronomy in the 20th Century*, Chapter XIV, Crowell Collier and Macmillan, New York, 1962.

Boyko, Anatole, "Inside a Globular Star Cluster," *Sky and Telescope*, p. 269, November 1964.

Brown, Harrison, "Planetary Systems and Main Sequence Stars," *Science,* p. 1177, September 11, 1964.
Kraft, Robert P., "Exploding Stars," *Scientific American,* p. 54, April 1962.
Limber, D. N., "The Pleiades," *Scientific American,* p. 58, November 1962.
Lippincott, S. and M. D. Worth, "The Double Star Sirius," *Sky and Telescope,* p. 4, January 1966.
Lovell, Sir Bernard, "Radio-emitting Flare Stars," *Scientific American,* p. 13, August 1964.
Saunders, J., "The Globular Cluster Omega Centauri," *Sky and Telescope,* p. 133, September 1963.
Struve, Otto, "Contact Binaries," *Sky and Telescope,* p. 197, April 1961.
Struve, Otto, "The Oldest Star Cluster," *Sky and Telescope,* p. 261, November 1962.
Van de Kamp, Peter, "Bernard's Star as an Astrometric Binary," *Sky and Telescope,* p. 8, July 1963.
Vandervoort, P. O., "The Age of the Orion Nebula," *Scientific American,* p. 91, February 1965.
Walker, Merle F., "New Observations of AE Aquarii," *Sky and Telescope,* p. 23, January 1965.
Worley, Charles E., "Measuring Double Stars with a Filar Micrometer," *Sky and Telescope,* p. 261, November 1961.
"An Ex-Variable Star?" *Sky and Telescope,* p. 323, June 1966.



The nebula and cluster M16. (Lick Observatory)

Chapter 14 *Between the Stars*

A good photograph of the Milky Way in the direction of the galactic center (in Sagittarius; see Figure 14-4) strikingly reveals the extent to which the stars near the galactic plane are immersed in gaseous clouds. But even such an excellent photograph does not reveal the gases and dust that are invisible though present in many regions of our galaxy. Observational methods more refined than those possible with direct photography have disclosed the existence of such interstellar matter. All interstellar matter can be divided into three main groups: the *bright nebulae,* the *dark nebulae,* and the gas and dust between the nebulae, generally called *interstellar gas* and *dust.*

Bright Nebulae

Bright nebulae (also called *diffuse nebulae*) are found only in close association with fairly luminous stars. It is these stars which cause such nebulae to become bright. They do so by one of two processes, which may be illustrated by a consideration of the Orion nebula (Figure 2-9) and the nebulosity in the Pleiades (Figure 13-12). A spectrogram of the Orion nebula discloses a bright-line spectrum, indicating that the gas in this nebula is at a low pressure and emits its own light. Such a nebula is therefore called an *emission nebula*. The spectrograms of the nebulosity about the Pleiades cluster, however, indicate that this nebulosity is merely reflecting the light of the embedded stars. The light from the nebulosity is recorded as an absorption spectrum similar to that of the brightest stars in the cluster. Such a nebula is therefore called a *reflection nebula.*

A pure gas will not reflect light, since the atoms composing such a gas are too small. Astronomers therefore assume that the light-reflecting agent in the reflection nebulae consists of small particles of "cosmic dust." The exact composition of this dust is not known with certainty, but we are able at least to estimate the size and nature of its particles by the light they scatter.

The light of the reflection nebulae is bluer than that of the stars whose light is scattered, although the difference in color is not as extreme as the difference between the red setting sun and the blue sky (see p. 70). The quality of this scattered light in the reflection nebulosity has led astronomers to believe that the dust particles causing the scattering are about 10^{-5} in. in diameter. (A simple molecule is about 10^{-8} in. in diameter.) Further studies have shown that the albedo of these dust particles is almost as high as that of white paper or snow. Such a high albedo indicates that they cannot be metallic but must be composed of hydrogen, carbon, nitrogen, oxygen, along with other elements, all perhaps combined chemically into molecules which join to form the dust particles.

The stars associated with reflection nebulae, although very luminous, are never of a spectral type earlier than B2. A nebula that has stars of spectral type B1 or earlier associated with it does not reflect very much light. Consequently the resulting nebular spectrum is not the same as the absorption spectrum of the brightest stars, but is instead a faint continuous spectrum upon which brighter emission lines are superimposed. These very hot luminous stars apparently reduce the number of light-scattering particles to a minimum either by blowing them away with radiation pressure or by evaporating them with their intense radiation.

Gases in an emission nebula emit a bright-line spectrum by means of a process called *fluorescence* (the same process that gives rise to commercial fluorescent lighting), which is activated by the very hot stars of B1 or earlier. Since these stars are so hot, they emit most of their light in the ultraviolet region of the spectrum. When the atoms of a material absorb energy of one wavelength (such as ultraviolet) and emit energy of a longer wavelength (such as visible light), the material is said to fluoresce. Ultraviolet radiation has more energy than visible light and thus is more capable of exciting the atoms of the gas. When an electron in an orbit around an atom is struck by a photon of ultraviolet light having just the right amount of energy (that is, just the right wavelength), the electron will absorb the photon and make a transition to a higher energy level within the atom. It will then fall back to the ground state and in so doing, it emits light (see p. 50). This light need not be of the same wavelength as the ultraviolet radiation absorbed, as the electron may return to the ground state by cascading down from one energy level to another until the lowest energy level is reached. In cascading the electron will emit photons of various but discrete amounts of energy, yielding the characteristic emission spectrum.

If a photon of ultraviolet radiation striking an orbital electron has sufficient energy, the electron will go free of the nucleus. This is the photoionization process discussed on p. 54, where we saw that any amount of energy above the minimum ionizing energy may be absorbed by the electron. Thus free electrons exist with different amounts of energy, and when one is captured it may give up, within set limits, any one of these amounts of energy.

The minimum amount of energy released by the capture of an electron falling to a given energy level is equal to the minimum amount of energy required to ionize the atom from that level in the first place. The maximum amount of energy given off in the capture of an electron by the nucleus is less definite and depends on whether or not the nucleus is able to capture the electron. If the electron has too much kinetic energy (that is, if it is traveling too fast) it will not be captured but will merely be deflected as it passes near the positive nucleus.

Since the wavelength of an emission line depends on the amount of energy released during the capture of an electron, and since the captured electron could have had any amount of energy within the stated limits, the many captures emit not one single spectral line but many immediately adjacent lines that blend to form a spectral continuum.

The spectral continuum caused by hydrogen atoms capturing electrons that fall to the second energy level (the Balmer level) is called the *Balmer continuum.* Captured electrons that fall to the first energy level correspondingly produce the *Lyman continuum* in the far-ultraviolet region. The Balmer continuum is in the near-ultraviolet region of the spectrum and thus is the more important for our discussion. It exists just to the short wavelength side of the line of shortest wavelength in the Balmer series of the hydrogen atom, which results from the transition of an electron from the highest energy level to the second energy level. If an electron is captured and then falls to the second energy level, however, it must release more energy than this transition because a captured electron will have more energy to give up than an electron falling from the highest energy level. Since a larger amount of energy results in a spectral line of shorter wavelength, the Balmer continuum is to the short wavelength side of the Balmer line series.

The Balmer continuum exists for absorption as well as for emission spectra. In absorption (as it is for the hotter stars), it results from photoionization. In emission it results from the capture of free electrons. Since each process involves ionization, the intensity of the Balmer continuum gives us some idea of the extent to which the gases are ionized. If the atoms are not ionized, the Balmer continuum cannot exist in either absorption or emission.

The intensity within the Balmer continuum also gives some idea of the kinetic energy of the free electrons. If the free electrons have a great deal of kinetic energy, the Balmer continuum will be more intense toward shorter wavelengths than if the free electrons had less kinetic energy. And since kinetic energy is directly related to temperature, the Balmer continuum gives us some idea of the temperature within the gases of a nebula. When the temperature is thus determined by the kinetic energy of the electrons it is called *electron temperature.* The electron temperature in many emission nebulae is estimated to be in the neighborhood of 10,000°K, which is substantially the temperature determined by studies of emission lines.

Emission lines in the spectra of gaseous nebulae, with the exception of the hydrogen lines, are not the usual spectral lines found in stellar spectroscopy. Although lines of nitrogen, oxygen, helium, and other elements are also found in nebulae, they result from transitions peculiar to emission nebulae. When the density of a gas is so extremely low as it is in gaseous nebulae, transitions occur that do not occur in the atoms of stars or in terrestrial laboratories. These transitions are so improbable in gases under pressures greater than those found in emission nebulae that they have been termed "forbidden transitions." Spectral lines resulting from such transitions are called *forbidden lines.*

Each electron transition is governed by what we call selection rules. These selection rules (derived in large part from the spectral line intensities), give the probability that any particular transition will occur. The most probable transitions result in the most intense lines in the spectrum of any given element. Transitions which are extremely improbable are called *forbidden transitions.* An undisturbed electron will remain in an energy level for a long time if to reach the ground state it must make a forbidden transition. Such an energy level is called a *metastable state* (meta-stable). An undisturbed electron in a metastable state may remain there for seconds, minutes, or even hours before making the forbidden transition to the ground state. Many such transitions result in a *forbidden spectral line.* An electron will remain in the more common energy level for only about 10^{-8} second. If an atom finds itself in a region of relatively high pressure such as the lower atmosphere of a star or a gas-filled tube in a college physics laboratory, it is being bombarded billions of times per second by the other atoms in the gas. Consequently its electrons cannot remain in metastable states long enough to make a transition down. The bombardment will knock an electron from the metastable state into a higher energy level from which it may make a downward transition within 10^{-8} second. For this reason forbidden lines are not usually seen in the terrestrial laboratory or in stellar spectra.

The gases in the nebulae, however, are under such low pressure that an atom may travel for days, weeks, or even years before it collides with another atom. If during this time one of its electrons falls into a metastable state, it may remain there long enough to make the downward transition that causes a forbidden line. Forbidden lines can result not only from neutral atoms but also from atoms that have been ionized once or twice. In emission nebulae forbidden lines result principally from ionized oxygen, nitrogen, neon, and sulfur.

It is interesting to compare the density of the Orion nebula with terrestrial densities. The Orion nebula has about 5,000 atoms (mostly ionized) per cubic in. One cubic in. of atmosphere at the Earth's surface contains about 4×10^{20} molecules. It is this lack of "congestion" in emission nebulae that creates the condition under which the atoms emit forbidden lines. However, the density of the Orion nebula is not the same throughout. In the brightest central part it is as "high" as 3.2×10^{5} (320,000) atoms per cubic in.

Despite this extremely low density, emission nebulae are visible. This is due to their tremendous size, which varies from nebula to nebula, although a diameter of 25 light years is fairly common. Thus, even though there may be only about 5,000 atoms per cubic in., an entire cloud contains many atoms, for there are about 5×10^{52} cubic in. in a cubic light year. To take our example, there are almost 7,000 cubic light years in the Orion nebula, which thus must comprise about 10^{60} atoms—enough material to make about 1,500 suns.

As was pointed out on p. 357, the visible phenomenon that we call the Orion nebula is only a small portion of a much larger gaseous cloud, a large part of which is invisible to direct photography. It thus appears that the photographic boundaries of the emission nebulae may not be their true limits but only their limits as defined by the extent to which the stars immersed in them excite the gases to emit their own light or reflect the light of those stars.

Dark Nebulae

The gaseous nebulae that neither reflect the light of embedded stars nor emit their own light are called dark nebulae. Their existence can be determined by the fact that they partially conceal luminous objects behind them. The Horsehead nebula in the Orion region is an excellent example (Figure 14-1). Not only does the dark gas outlining a horse's head conceal the emission nebula behind it, but by comparing the number of stars on the left of the photograph with the number on the right it can be seen that the Horsehead nebula is only a projecting branch of a much larger dark nebula that conceals the more distant stars on the left.

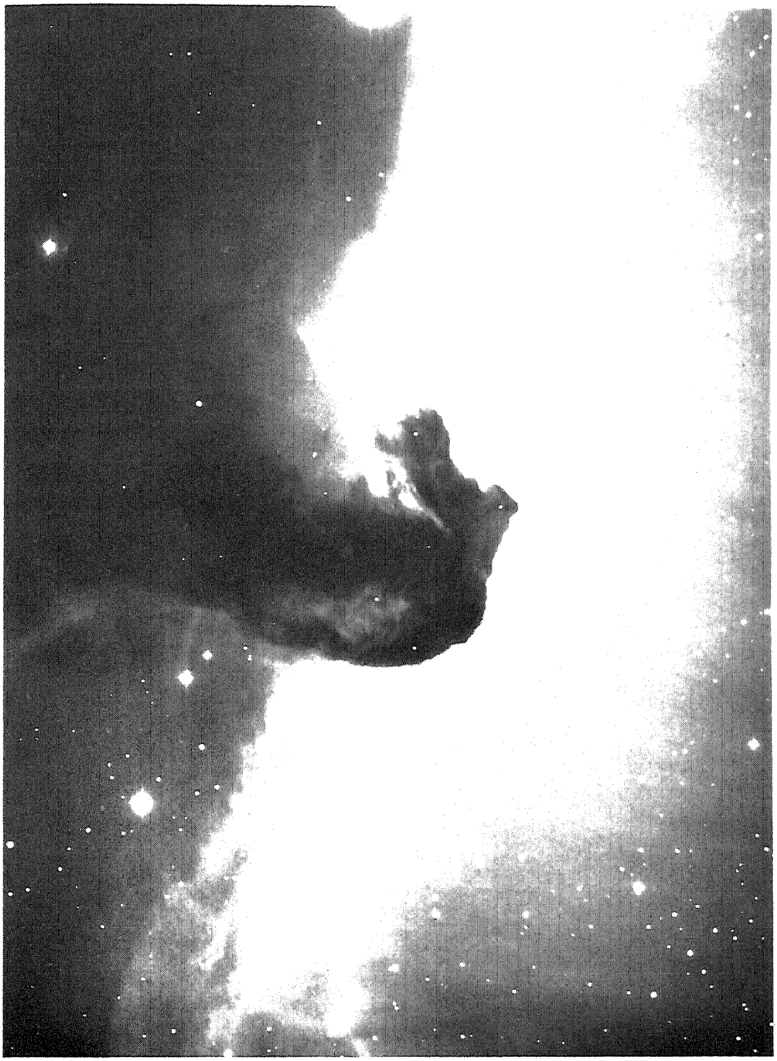


Figure 14-1 **The Horsehead nebula in Orion. (Mount Wilson and Palomar Observatories)**

Evidence of the coexistence of dark and bright nebulae in the same region is also seen in Figure 14-2, a photograph in the region of the star Rho Ophiuchi. Lanes of dark nebulosity can be seen extending from the central portion of the photograph where gases partially conceal stars behind them. In this nebula the bright portion is largely reflection nebulosity.

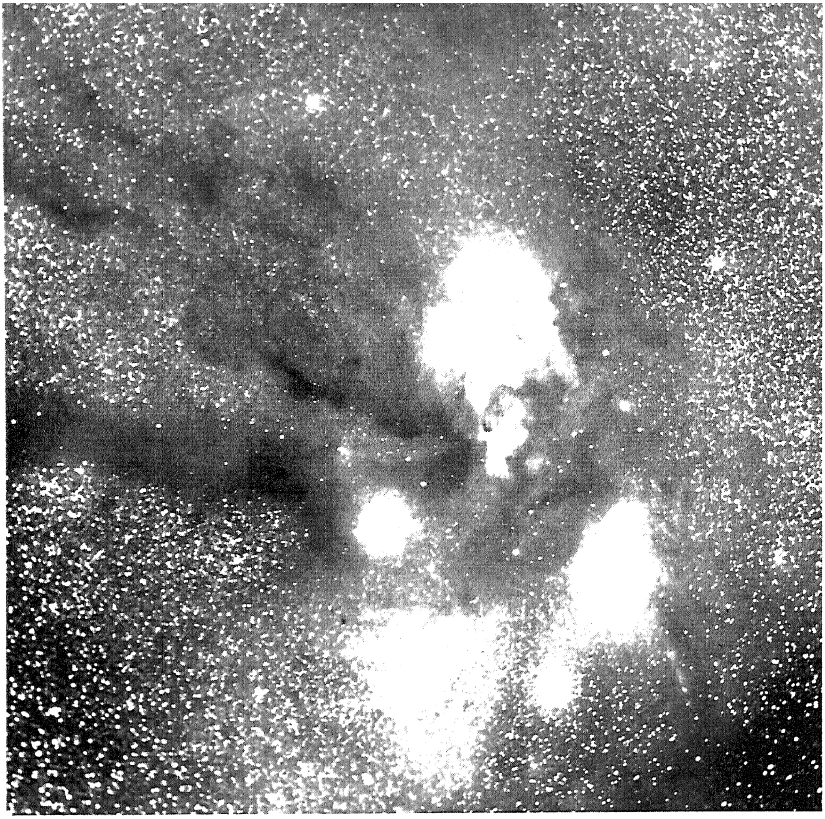


Figure 14-2 Close association of bright and dark nebulae near the star Rho Ophiuchi. Two globular clusters can be seen just below the center of the bright nebulosity. (Yerkes Observatory)

There are many other dark nebulae appearing as "voids" in space that are actually large masses of gas and dust with no bright stars nearby to illuminate them. The Coalsack in the southern hemisphere is an excellent example (Figure 14-3). A photograph of the star "clouds" in Sagittarius (Figure 14-4) not only tells us that our sun is not alone (anyone for star counting?) but also tells us that between the Earth and this prodigious number of stars intervene concealing dark clouds. The true extent of the dark nebulae can best be seen, however, by a contemplative look at the Milky Way when one is free of the stifling effect cities have on astronomical observing. Run-

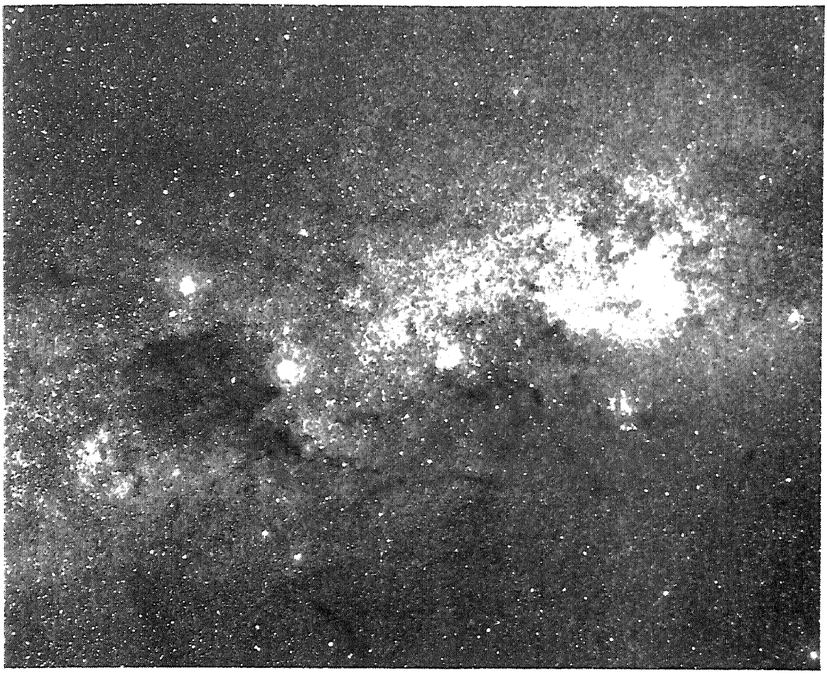


Figure 14-3 The Coalsack can be seen easily with the naked eye in the southern Milky Way. Just above and to the right of the Coalsack are the four stars of the Southern Cross. Bright nebulosity can be seen to the right. (Harvard College Observatory)

ning nearly the entire length of the Milky Way is the "Great Rift" which seems to divide the Milky Way into two parts. This is not a real rift in the groupings of stars, but an apparent rift caused by intervening gas and dust.

The sizes and distances of the dark nebulae present a new problem, for all the methods of distance determination discussed so far rely on the fact that objects emit light. The dark nebulae do not. We can, however, fall back on the stars to give us some idea of the distances and thus of the sizes of dark nebulae.

It is possible to estimate the distance of a dark nebula by comparing the number of stars of each apparent magnitude in the region of the dark nebula with a nearby region which appears to be unobscured. In general, the number of stars increases with decreasing brightness. The rate of increase of the number of stars which lie between us and the dark nebula will be the same as the rate of increase for

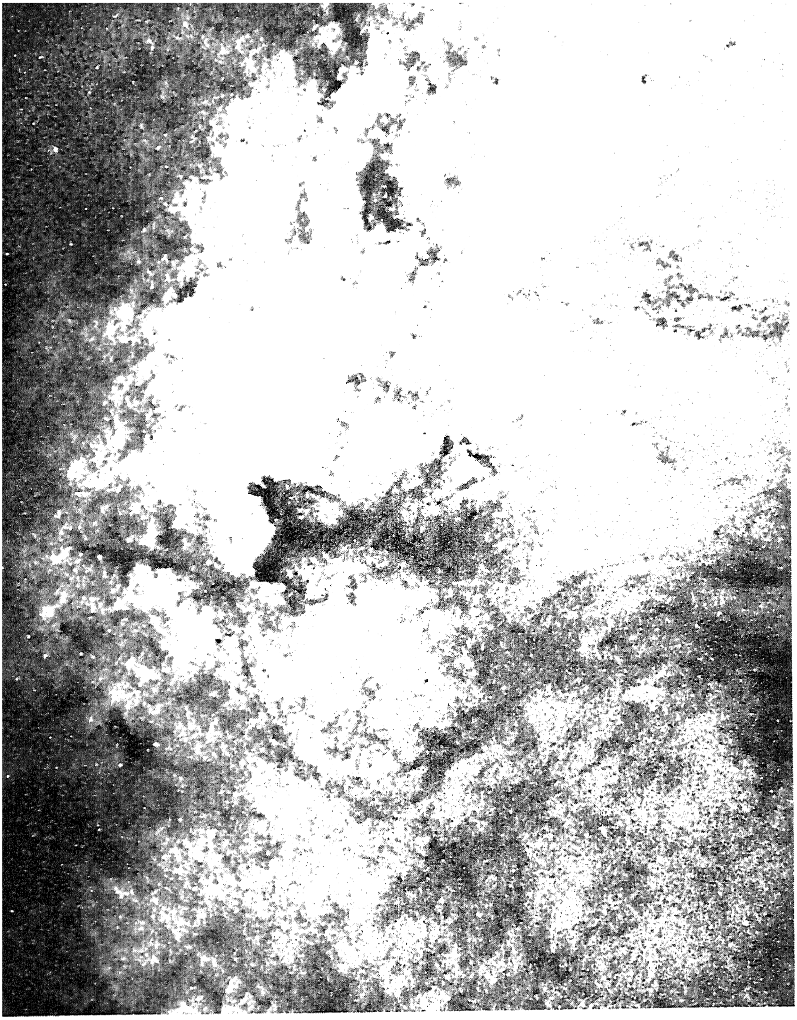


Figure 14-4 Dark clouds in the region of Sagittarius photographed in red light with the 48-in. Schmidt telescope. (Mount Wilson and Palomar Observatories)

the unobscured region. This is illustrated in Figure 14-5 as the line which extends unbroken to about the eleventh magnitude. For stars fainter than the eleventh magnitude, however, the rate of increase for the region of the nebula (the dashed line) is less than the rate of increase in the unobscured region. The reason, evidently, is the dimming effect of the nebula; and, if we can estimate on a statistical

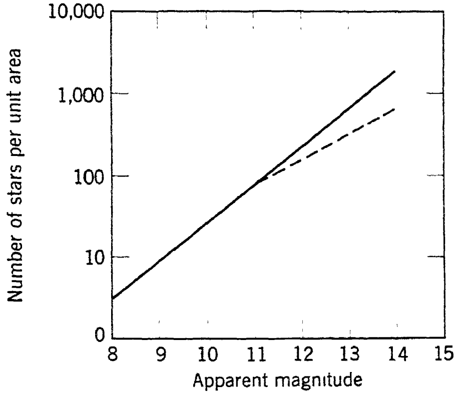


Figure 14-5 The distance of dark nebulae can be determined by comparing star counts in adjacent obscured and unobscured regions.

basis the average absolute magnitude of the stars involved, we can estimate the distance of the nebula. By knowing the distance and the angular diameter of the dark nebula, we can calculate the linear diameter by the use of trigonometry.

From studies of this kind the Coalsack is estimated to be only 400 light years away and to have a diameter of roughly 40 light years. Since it can be seen easily with the naked eye, it is interesting to speculate how it would appear if there were a cluster of O and B stars immersed in it.

Dark nebulae seem to be composed of about the same material as bright nebulae. The stars whose light shines through the dark clouds are reddened to such an extent by having their light scattered that it appears likely that the dust particles are about the same size as those in the reflection nebulae.

Interstellar Gases and Dust

Now that we recognize the existence of dark gases, how can we be sure that the unobscured regions are really unobscured? If there is material between the stars that does not reveal itself as obvious dark nebulae, we might be able to detect it by the effect it would have on the light passing through it. There are two possibilities. If the material were to contain dust particles, the light could become slightly reddened by scattering; or if the material were to contain gas, it could cause additional absorption lines in the spectra of the stars whose light shines through the material.

To find evidence for this interstellar material we should study stars that we know are at large distances from the Earth, for if such an effect is detectable it is more likely to be noticeable in distant stars whose light would have to travel through more interstellar material. Furthermore, the stars we study should lie near the galactic plane, where the vast majority of nebulae appear; for example, we could observe the faint B-type stars. B-type stars have a high luminosity; consequently, if they appear faint in a region of little obscuration they must be far from the sun.

Such studies do in fact reveal the existence of both interstellar gases and dust. The faint B stars are considerably redder than they should be. The amount of reddening can be measured by comparing their observed color index, as determined by photoelectric studies, with their expected color index, as determined from their spectral type. The difference between the two color indices is called the *color excess*. This term is derived from the fact that the color index increases in a positive sense for increasingly cool stars. Thus stars reddened by interstellar dust have greater color indices than they would normally have (see p. 353f). Color excess is a misleading term, however, since the increase in the color index is really a measure of the amount of blue light that has been *subtracted* from the star's light by scattering. The term is thus descriptive of no more than the reddening of the light that reaches us.

The existence of interstellar gas (but not dust) can be verified by analyzing the spectra of the faint B stars. The spectra of most of them show a few very sharp, although faint, lines that result from interstellar gases. These *interstellar lines* were first noted in the spectra of spectroscopic binaries because such lines do not take part in the periodic oscillation of the stellar lines (Figure 13-8).

The more common interstellar lines are the H and K lines of ionized calcium and lines from neutral calcium, potassium, iron, sodium, and ionized titanium. However, interstellar lines are not restricted to these elements, for interstellar molecular bands also appear. Some of these bands have not yet been identified, although bands of CH, CN, and CH^+ have been identified. The molecule OH has been definitely identified with radio telescopes. The OH molecule produces spectral lines in the radio region of the spectrum at frequencies of 1720, 1667, 1665, and 1612 megacycles (wavelengths between 17.4 and 18.6 cm), and with an expected intensity ratio of 1:9:5:1 respectively. Radio spectral lines in both emission and absorption at these frequencies but not with this ratio of intensity have indeed been discovered and studied by groups in Massachusetts, California, Australia, and England. The fact that their intensity ratio differs from the straightforward prediction indicates that some action, as yet unaccounted for, is altering the expected probability of transitions.

In some stars, more than one interstellar line from a given transition in an element will appear, giving evidence that the interstellar gases are not uniformly distributed, but form numerous clouds with perhaps even more tenuous gas and dust between these clouds. Each cloud absorbs light independently of the others, and if each has a different velocity with respect to the Earth, it will produce an absorption line with a different Doppler shift from the others. In at least four stars the interstellar K line of calcium appears four times, indicating four different intervening clouds. One star has as many as seven components of the same interstellar line.

The density of interstellar gases is so low that by comparison even the emission nebulae seem almost congested. It has been estimated that in a cloud of interstellar gas there are on the average 200 atoms per cubic in., but between these clouds we may find only 20 atoms per cubic in. Here, then, is an opportunity for high-vacuum studies. Indeed, astronomical observations serve to supplement and extend the observations made in the physics laboratory, for in the celestial realm we find and study atoms at temperature and pressure extremes which we could not hope to obtain in terrestrial laboratories.

The composition of all interstellar gases, including diffuse nebulae, does not seem to differ much from what we find in the sun, in most of the stars, and, with the exception of hydrogen and helium, from what we find here on the Earth. The Earth as well as Mercury, Venus, the moon, and Mars do seem unusual in their comparative lack of hydrogen and helium.

There is evidence to support the idea that a good part of the interstellar dust is particles of carbon, perhaps like graphite. The main evidence is the manner in which interstellar dust reddens the light which passes through it. This reddening means that with decreasing wavelength less light passes through the dusty clouds; with decreasing wavelength more light is scattered out of the beam. Reddening of the same quantitative nature has been produced by graphite grains in the laboratory. Besides, carbon is abundant enough in the universe to supply the interstellar regions with the amount of dust we observe.

Thus our galaxy not only consists of some 100 billion stars but also contains large amounts of gas and dust between them. This gas and dust is not distributed uniformly throughout the galaxy but appears in the form of individual clouds close to the galactic plane. If the density of a cloud is great enough, it will partially obscure the stars behind it and be called a dark nebula. If there are enough stars in the cloud, it will become a bright nebula. If those stars are B2 or later, the nebula will shine by reflected light and be called a reflection nebula. If the stars are B1 or earlier, the dust tends to be forced away and the nebula will shine mainly by fluorescence, having absorbed ultraviolet radiation from the hot stars and emitting the visible light that makes it an emission nebula. If the gas and dust have a density too low to be readily noticeable as a dark nebula, they are considered as interstellar material whose presence is detected by the scattering of the blue light from the stars that shine through the nebula or by the interstellar absorption lines produced in the spectra of those stars.

Radio Astronomy

There is so much interstellar material that a major part of our galaxy remains concealed from visual observation, and the nucleus of our galaxy is hidden from our view. But the advent of radio astronomy has given us a tool for penetrating interstellar clouds because scattering decreases with increasing wavelength. Thus radio waves, with wavelengths a million or more times longer than that of light, travel through interstellar material without appreciable absorption or scattering.

Since radio waves are part of the electromagnetic spectrum, we can speak in terms of a continuous radio spectrum as well as of an emission or absorption radio spectrum. An emission radio line is equivalent to an emission line in the visible spectrum. There is also the possibility that a gas could selectively absorb radio radiation and produce the equivalent of an absorption line. Let us first consider the continuous spectrum in the radio region.

One of the radio astronomers' first tasks was to determine where in the sky the radio signals were coming from; in effect they asked, "What does the sky look like when 'seen' by radio waves?" The best way to answer such a question is to make a survey of the sky by

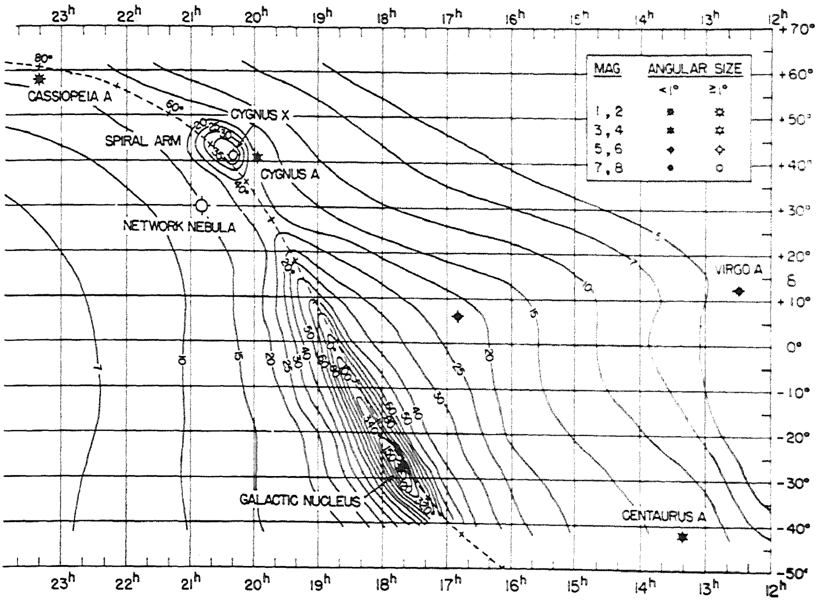


Figure 14-6 Radio map of the sky made at a wavelength of 120 cm with the Ohio State University radio telescope. (Radio Observatory, Ohio State University)

recording the intensity of the radio radiation from every part of the celestial sphere. The results of such a survey look like a contour map. On a geographical contour map the contour lines connect points of equal elevations, whereas on a radio map of the sky the contour lines connect points of equal radio radiation intensity (Figure 14-6).

The results of many radio maps of the celestial sphere made at various wavelengths indicate that: (*a*) the stars are very weak emitters of radio energy, (*b*) the Milky Way, with all of its gas, is a major source of radio energy, and (*c*) there are many *discrete radio sources* in the sky.

No star is known to appear on this or any other radio map. Stars simply do not radiate much energy in the region of the radio or microwaves. If our sun is a powerful radio source, it is because of its proximity to us. There is, however, one exception to this: *flare stars* have been observed simultaneously with both optical and radio telescopes. But they can be observed by a radio telescope only during the flare. Flare stars are dwarf stars of spectral type M5 or M6 with emission lines in their spectra. Since flare stars have been detected in the Pleiades, it is not unreasonable to suspect that they may be young stars. But a number of stars in the solar neighborhood are well-established flare stars and are not members of a young cluster.

Simultaneous observations of flare stars have been achieved in both England and Australia. The optical observations were made by the Baker-Nunn satellite tracking cameras operated for the NASA, and by competent amateur astronomers. Radio observations were made with the 250-ft radio telescope at Jodrell Banks, England, and with an interferometer-type radio telescope in Fleurs, Australia. It seems clear that the flare star emits more energy in each flare than does the sun during a typical solar flare. Therefore it would seem as if the two events are not of the same origin. But observations have only begun on this interesting development, so there is much yet to learn.

The contour lines in the radio map of Figure 14-6 represent the amount of energy received by the radio telescope. The numbers of each contour line refer to particular units of energy. By comparing these numbers on the map it can be seen that in general the amount of energy increases toward the central plane of the Milky Way. In the direction of the galactic nucleus, the amount of energy received reaches a high of 150 units.

Energy from the gases between the stars originates in one of several ways. If the gas is largely ionized, it is composed mostly of free electrons and free protons. On occasions a free electron travels close to a free proton, but travels too fast to be captured. Nevertheless its line of travel is deviated by the electric attraction of the two particles. During the change in the direction of the travel the electron emits energy; a change in direction of travel amounts to a transition from one "free orbit" to another "free orbit." Such a transition is called a *free-free transition*. Since these free orbits are not quantized, the free-free transitions may result in the emission of any amount of energy within certain limits. The limits are such that this energy is generally emitted in the radio region of the electromagnetic spectrum. And since the energy of the free orbits is not quantized, the radio energy emitted amounts to a continuous spectrum—all energies are emitted. If the intensity of each wavelength corresponds to Planck's law, then we surmise that the source is thermal in nature and the gas is largely ionized. The possibility exists that the continuous radio spectrum is synchrotron radiation, in which case it is polar-

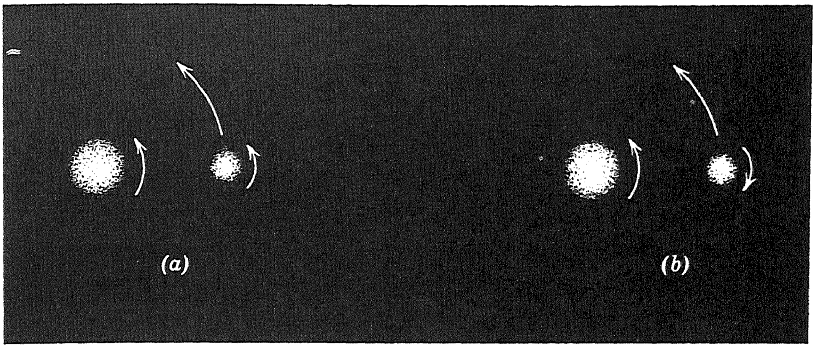


Figure 14-7 The 21-cm line in emission is caused by the electron spin flipping from (*a*) to (*b*).

ized and its energy distribution is different from that of a thermal source (see p. 290ff).

As was pointed out earlier in this chapter, radio sources can also emit energy in a form equivalent to an emission line. One such radio emission line has a wavelength of 21 cm (about $8\frac{1}{3}$ in.). This line is caused by a transition of an electron within the hydrogen atom. This particular transition, however, is slightly different from those discussed so far.

An orbital electron in an atom spins on an axis, like the Earth, as it revolves about the nucleus. But the nucleus also spins, like the sun. If the *electron spin* is in the same direction as the *nuclear spin,* the atom has slightly more energy than if the two spins are in opposite directions. Thus if they change from spinning in the same direction to spinning in the opposite direction, energy will be released (Figure 14-7). The amount of energy released is very small but is always the same and thus gives rise to an emission line. If the electron is in the ground state of a hydrogen atom when they reverse their spin, the amount of energy released or absorbed corresponds to a wavelength of 21 cm, or a frequency of 1420 megacycles.

The existence of this emission line was predicted in 1945 by the Dutch astronomer Hendrik C. van de Hulst, but the chances of detecting it looked rather poor at first, since the probability of a given hydrogen atom's making such a transition is extremely low. Any given hydrogen atom, if undisturbed, should make this transition once every 11 million years!

But the hydrogen atoms in interstellar space are not left completely undisturbed. An interstellar hydrogen atom will collide with one of its distant neighbors (less than a centimeter away) once about every 300 years! Even though these collisions are very infrequent as compared with the collisions of molecules in the Earth's atmosphere (which occur at the rate of more than a billion times each second) they are frequent enough to increase the probability of a hydrogen atom's changing its electron spin. This increased probability, coupled with the fact that there are billions upon billions of hydrogen atoms in interstellar space, led radio astronomers to believe that the 21-cm line could be found.

The 21-cm line was finally detected in 1951. Since then it has assumed a major role in the study of our galaxy, for with it astronomers can detect clouds of neutral hydrogen from nearly the entire galaxy. The 21-cm line can only be emitted by a neutral hydrogen atom since it is a transition within that atom. The clouds of neutral hydrogen presumably contain ionized hydrogen as well, which emits a continuous spectrum of radio energy on which the 21-cm line is superimposed.

As we have seen, the interstellar clouds of gas and dust do not scatter the 21-cm line as they do visible light. Thus this radiation can travel through regions of space opaque to visible light. We should note, however, that the 21-cm line has also been detected in absorption which indicates that its energy can cause a hydrogen atom to change its electron spin back to the higher energy level where both electron and nucleus spin in the same direction.

The 21-cm line of hydrogen and the OH spectral lines offer an additional advantage not offered by the continuous radio emission: they can be used to detect radial velocities of clouds of neutral hydrogen by their Doppler shift. The intensive studies made with the 21-cm line, as well as with continuous radio emission, have given us a much better idea of the structure of our galaxy than we had at mid-century.

Of the many discrete radio sources some of them are the gaseous nebulae. Emission nebulae generally emit continuous energy indicating that the gases are largely ionized. Further studies have determined a temperature of about 10,000°K in agreement with optical measurements. Furthermore, radio studies lead us to believe that if the exciting star is A0 the gas is ionized out to about 3 light years. However, if the exciting star is O5, the gases will be ionized out to as far as 490 light years. On the average, the electron density in these clouds is about 10 to 50 electrons per cubic cm increasing to a few thousand in the central regions. In each of the nebulae studied (including the Orion nebula, the Rosette nebula, the Lagoon nebula) the radio radiation was thermal in origin.

However, the problem of telescopic resolution faces us again. If a discrete radio source is a point source, then radio interferometers come to our aid and can be used to establish its position. But if the discrete radio source is extended, such as the nebulae, then interferometers cannot be used to full advantage. Those sources which lie close enough to the ecliptic may be occulted by the moon. As the moon passes in front of and covers the radio source the signal we receive tells us something of the energy distribution in the source; then, as the moon moves on and uncovers the source we obtain still more information. Not all of the astronomer's observational techniques are simple observing. The astronomer learned long ago to use his ingenuity.

Basic Vocabulary for Subsequent Reading

Balmer continuum
Color excess
Diffuse nebula
Discrete radio source
Emission nebula
Fluorescence
Forbidden line
Metastable state
Reflection nebula

For Further Reading

Steinberg, Jean Louis, and James Lequeux, *Radio Astronomy*, McGraw-Hill Book Co., New York, 1963.

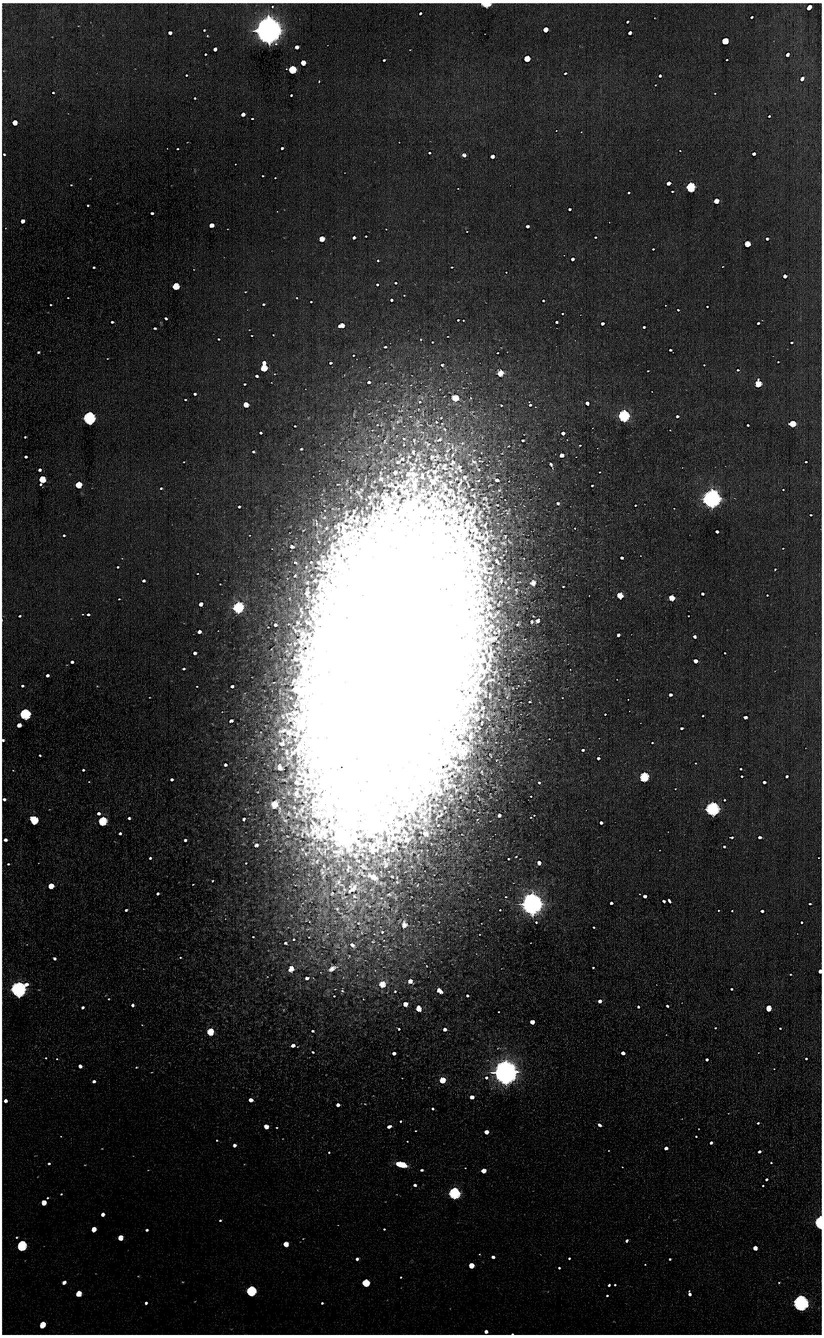
Struve, Otto, and V. Zebergs, *Astronomy in the 20th Century*, Chapters XVII, XVIII, Crowell Collier and Macmillan, New York, 1962.

Bowyer, C. Stuart, "Galactic X-ray Astronomy," *Sky and Telescope*, p. 264, November 1965.

Clarke, R. W., "Locating Radio Sources with the Moon," *Scientific American*, p. 30, June 1966.

Dieter, Nannielou, Harold Weaver, and David R. W. Williams, "The Interstellar Hydroxyl Radio Emission," *Sky and Telescope*, p. 132, March 1966.

Friedman, Herbert, "X-ray Astronomy," *Scientific American*, p. 36, June 1964.
Lilley, A. E., "The Absorption of Radio Waves in Space," *Scientific American*, p. 48, July 1957.
Robinson, Brian J., "Hydroxyl Radicals in Space," *Scientific American*, p. 26, July 1965.
Struve, Otto, "Interstellar Gas Clouds," *Sky and Telescope*, p. 269, May 1961.
"Graphite Grains in Interstellar Space," *Sky and Telescope*, p. 10, July 1963.



Elliptical galaxy—a satellite of the Andromeda galaxy. (Mount Wilson and Palomar Observatories)

Chapter 15 *Galaxies*

Our intimate relation with our own galaxy is both an advantage and a disadvantage to the astronomer who wishes to learn more about its structure, rotation, and evolution. In Chapter 14 it is pointed out that since we are near the central plane of the galaxy and therefore in nearly the same plane as most of the interstellar material, much of the galaxy is hidden from visual observations. Radio observations, of course, tend to alleviate this disadvantage; nevertheless we would still like to *see* what our galaxy looks like. The problem is not unlike that of placing your head between the spokes of a large wagon wheel and then trying to figure out what the rest of the wheel looks like. Without previous knowledge of wagon wheels in general, the problem would be a difficult one. Correspondingly, a knowledge of galaxies in general is useful to the astronomer. Astronomers therefore study other galaxies in great detail, for not only do these studies help us interpret observations of our own galaxy, but they also add to our knowledge of how the universe is put together. Let us investigate the structure of galaxies in general before we consider our own in detail.

Galaxies appear in every portion of the sky except along the Milky Way where the dust and gas of our own galaxy conceal the galaxies beyond. Hundreds of millions of galaxies are seen in the universe; the most distant ones visible in our large telescopes are estimated to be between 5 and 10 billion light years away. But galactic structure can be studied in only those that are much closer to us. The late Edwin P. Hubble, of the Mount Wilson Observatory, made an intensive study of galaxies and recognized three basic structures for the nearer galaxies, *elliptical*, *spiral*, and *irregular*. Each basic structure is further subdivided according to variations of shape within that group.

The elliptical galaxies (frontispiece Chapter 15) seem to be similar in appearance to the globular clusters although they are much larger and contain many more stars. They vary in shape from spheres to ellipsoids whose major axis is about five times longer than the minor axis. Since rotation causes flattening, the faster a galaxy rotates the more ellipsoidal it is.

Hubble's classification of elliptical galaxies is shown in Figure 15-1; the letter E (standing for elliptical) is followed by a number that indicates the degree of ellipticity. An E0 galaxy presents a circular disk. With increasing ellipticity the number following the E increases to an arbitrary maximum of 7. Since it is very difficult to determine the orientation in space of an elliptical galaxy, this classification applies to its shape only as it is seen from the Earth.

The spiral galaxies present a dynamic picture (Figure 15-2) that stands in contrast to the rather uniform elliptical galaxies. The first impression received upon looking at a spiral galaxy is that it is rotating about a central nucleus from which *spiral arms* curve out into space. Spiral galaxies were classified by Hubble into three subgroups: Sa, Sb, and Sc, according to the amount of material in the arms relative to that of the nucleus, as well as to the degree of openness of the arms. Examples of this classification are seen in Figure 15-2*a*, under the heading of *normal spirals*.

In the same group with the normal spirals are the *barred spirals* (Figure 15-2*b*), so called because they have a "bar" of stars running

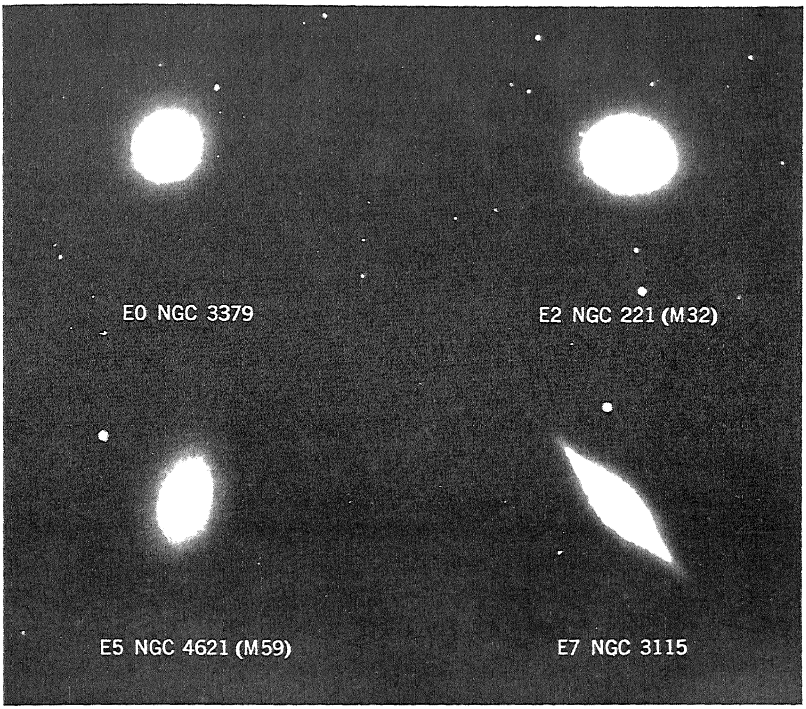


Figure 15-1 The shapes of elliptical galaxies vary from spherical to ellipsoidal. (Mount Wilson and Palomar Observatories)

through the nucleus. From the ends of this bar spiral arms may form, sometimes into a ring about the nucleus. Why some spiral galaxies assume the shape of normal and others of barred spirals is not known. The normal spirals, however, are more numerous: of the 600 galaxies studied by Hubble, 17% are elliptical; 50% are normal spirals; 30% are barred spirals; 3% are irregular.

The third type, the irregular galaxy, has no definite form. The two Magellanic Clouds are often considered to be members of this group, although some astronomers believe that the Large Magellanic Cloud should be classified as a barred spiral.

Elliptical Galaxies

It was not until the late 1940's that new photographic techniques made it possible to resolve one of the nearest elliptical galaxies, a companion of the Andromeda galaxy, into separate stars (frontispiece Chapter 15). The distribution of stars in elliptical galaxies is such that

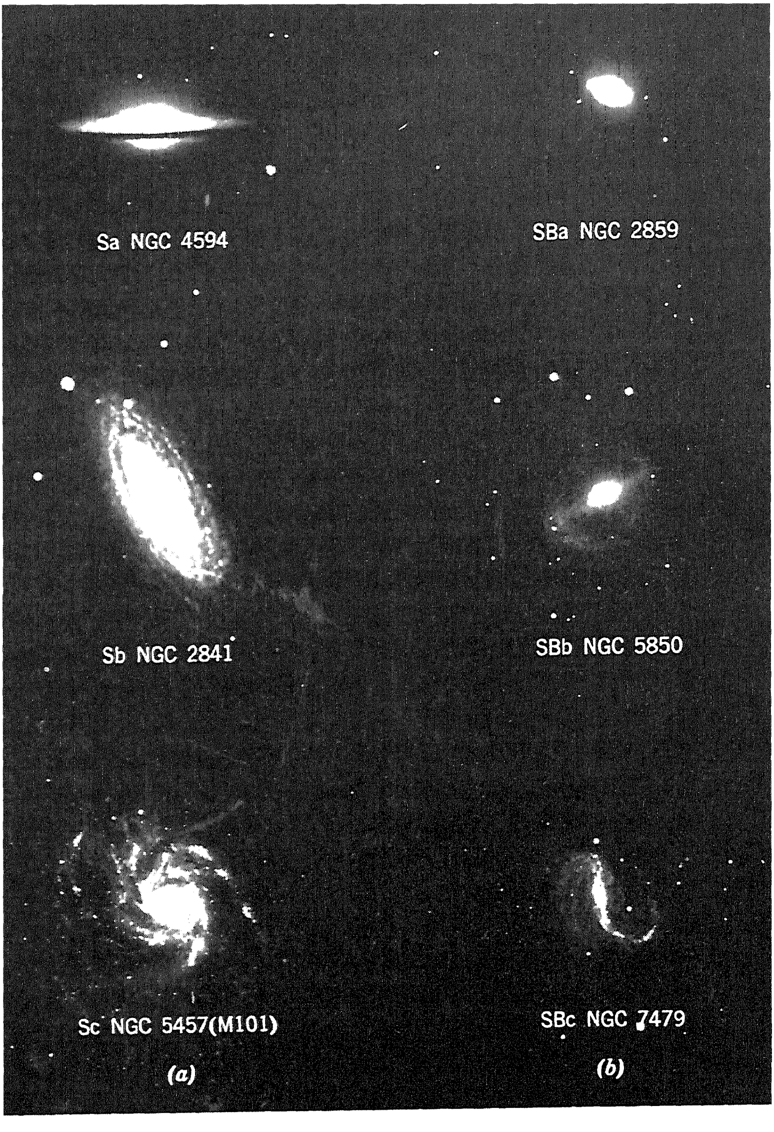


Figure 15-2 Different types of (*a*) normal spiral galaxies and (*b*) barred spiral galaxies. (Mount Wilson and Palomar Observatories)

the over-all shape of the galaxy does not change on photographs taken with different time exposures. A shorter exposure records only the more densely populated central regions, whereas a longer exposure records the outer regions of the galaxy. Thus is appears that the stars are placed rather symmetrically about the center but decrease in number from the center out.

The similarities of an elliptical galaxy, the nucleus of a spiral galaxy, and a globular cluster are striking. The brightest stars in each are red supergiants; the blue giants are conspicuous by their absence. It is not surprising, therefore, to learn that elliptical galaxies also contain RR Lyrae stars and are practically devoid of interstellar material. The elliptical galaxies are therefore classified as Population II. Since the lack of interstellar material makes them nearly transparent, more distant galaxies can sometimes be seen through their outer regions. It has recently become clear that the elliptical galaxies are the largest and most massive type of galaxy.

Normal Spiral Galaxies

Spiral galaxies differ considerably from elliptical galaxies both in structure and in stellar content. The nucleus, of course, resembles an elliptical galaxy, but here the similarity ends. The spiral arms contain many blue-giant stars, open clusters, and gaseous nebulae. The Andromeda galaxy, M 31 (briefly discussed on p. 281), is typical of the normal spirals and is close enough to our own galaxy (2.2 million light years from the sun) to be faintly visible to the naked eye as a hazy patch in the sky. Moreover, in a photograph the details of the Andromeda galaxy can be clearly seen (Figure 15-3). The central portion is the nucleus composed of Population II stars. The spiral arms can be traced by following the lanes of emission nebulae, dark gases, galactic clusters and blue giants—all Population I objects.

Despite the brightness of the spiral arms, most of the light from the region outside the nucleus comes from a substratum of stars whose luminosity is equal to or less than that of giants. The arms are superimposed upon this substratum, which is nearly transparent, since more distant galaxies are visible through the parts of it that lie between the spiral arms. Observations indicate that the shape of the substratum is somewhat ellipsoidal. It extends above and below the plane of the galaxy, includes many globular clusters, is redder in color than are the spiral arms, and thus may be Population II.

Since we see the Andromeda galaxy at an oblique angle (the plane

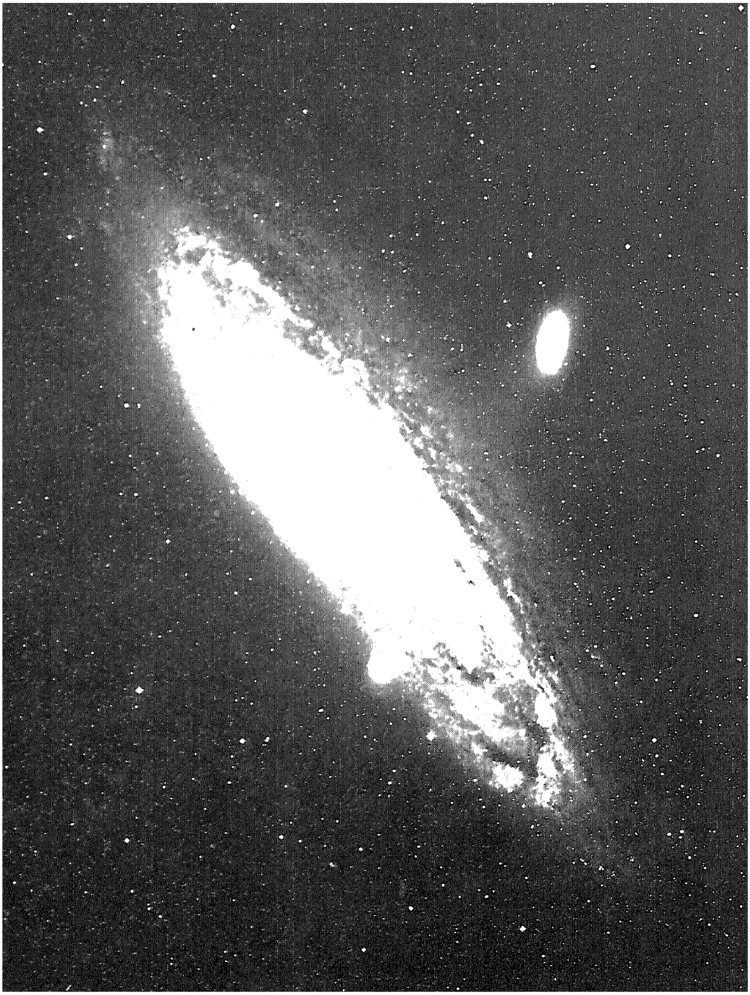


Figure 15-3 The Andromeda galaxy. (Mount Wilson and Palomar Observatories)

of the galaxy makes an angle of about 12° with the line of sight), we are able to determine by the Doppler shift the rate of rotation of those parts that are either approaching or receding from us. N. U. Mayall, of the Kitt Peak National Observatory, has studied the rotation by observing emission nebulae in the spiral arms. The emission nebulae give sharper spectral lines than the stars do and are therefore better suited for radial velocity measures. Mayall's results

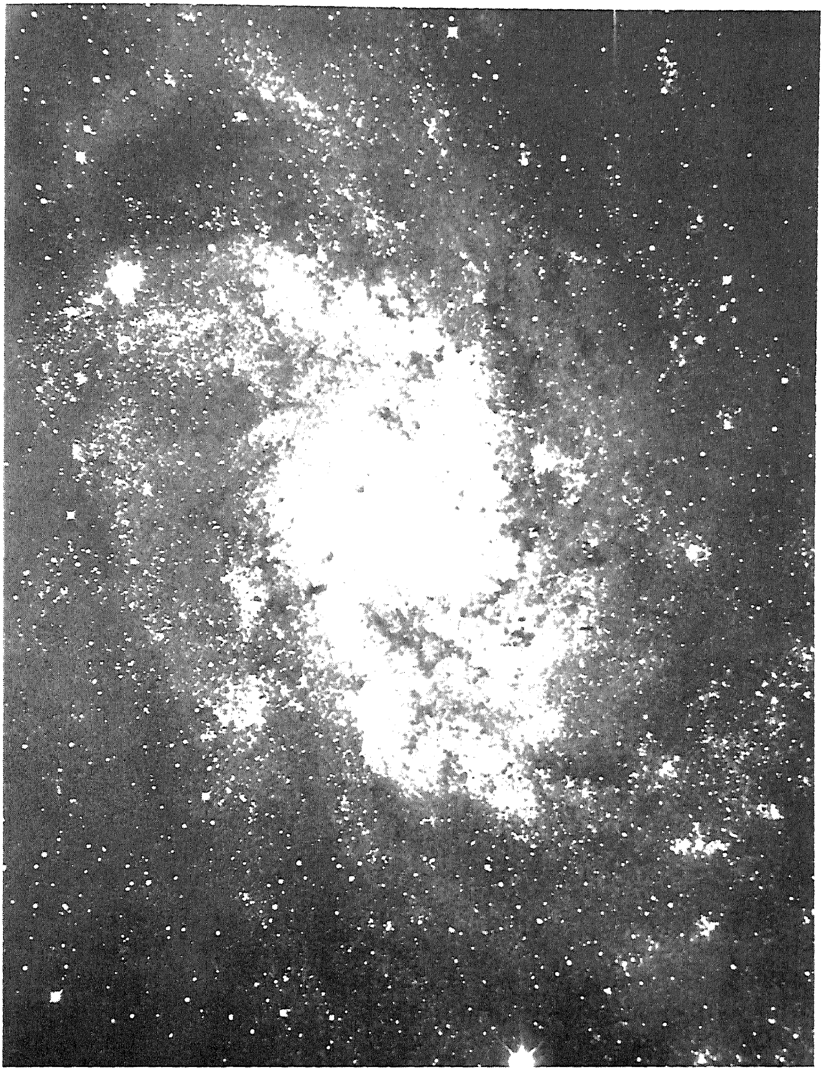


Figure 15-4 The spiral galaxy in Triangulum, M 33. (Mount Wilson and Palomar Observatories)

show that the inner portion of the galaxy rotates like a solid wheel; that is, the speed of the revolving stars increases with increasing distance from the center of the galaxy. The outer parts, however, rotate in a manner not unlike the motion of the planets about the sun; that is, the speed of the revolving stars decreases with increasing distance from the center.

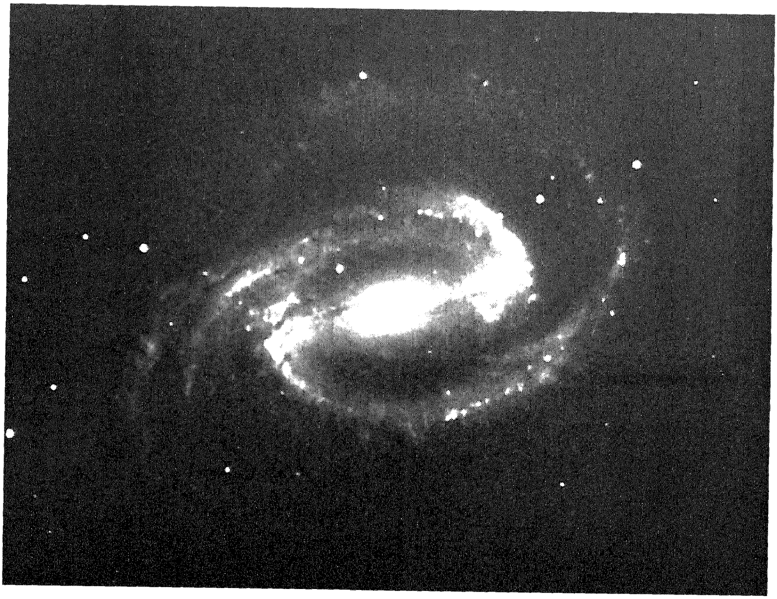


Figure 15-5 Barred spiral galaxy NGC 1300. (Mount Wilson and Palomar Observatories)

If the vast majority of the mass of any system is at the center, we can expect the system to rotate in accordance with Kepler's third law. If, however, the mass of the galaxy is distributed with uniform density throughout the entire galaxy, we can expect it to rotate like a solid wheel. The Andromeda galaxy lies somewhere between these two extremes.

Since we know the distance of the Andromeda galaxy, we can determine its size from its angular diameter. The portion that appears in Figure 15-3 has a diameter of about 100,000 light years. Measurements made with a photocell, however, indicate that there is a fringe of stars extending beyond the obvious photographic boundaries. The diameter including this fringe is about 180,000 light years. From the size and rotation of the Andromeda galaxy, the mass of this system has been estimated to be 4×10^{11} solar masses. Since the sun is roughly an average star in mass, we can assume that there are some 400 billion stars in the Andromeda galaxy!

Barred Spiral Galaxies

Barred spirals are similar to normal spirals in that their nuclei contain principally Population II objects and their arms are composed of Population I objects. Bright nebulosities as well as open clusters and lanes of dark obscuring matter can be seen in the spiral arms. The significant difference between normal and barred spirals can be seen by comparing Figure 15-4 with 15-5. The bar and the nucleus seem to be rotating as a unit, like two spokes extending from the hub of a wheel. The spiral arms seem to lag behind the bar and the nucleus; thus they appear to rotate in a manner similar to that in which our planets revolve. At present there is no explanation for the structure of barred spirals. However, it is suspected that magnetic fields play a major role in the structuring of galaxies including the barred spirals.

Irregular and Dwarf Galaxies

The irregular galaxies are quite different from both the elliptical and the spiral galaxies. They have no central nucleus, no spiral arms, and apparently no plane of symmetry. Common examples of this type of galaxy are the Magellanic Clouds (Figure 15-6).

The two Magellanic Clouds are located fairly close to our galaxy, at a distance of only about 150,000 light years. The diameter of the main body of the Large Cloud is 20,000 light years, but including its out-lying clusters its diameter is about 40,000 light years. The over-all diameter of the Small Cloud is about 25,000 light years. The mass of the Large Cloud is about 0.1 times that of our own galaxy, and the luminosities of each cloud seem to be the average for galaxies in general.

The Large Magellanic Cloud is a most interesting object for astronomers concerned with the evolution of stellar systems. It has an abundance of Population I objects: blue-giant stars, open clusters, obscuring gas and dust, and emission nebulae. In fact, the largest emission nebula known, the Tarantula nebula, lies in the Large Cloud. It is so big that were it to replace the Orion nebula, it would cover the entire constellation of Orion and be an extremely obvious feature in our night skies. But the most curious objects in the Large Cloud are the globular clusters, for some of these globular clusters have blue-giant stars! These globular clusters must be young and some may still be in the process of formation. There are a few clusters which are very regular and globular in shape but their color-luminosity diagrams look like those of the open clusters in our galaxy. These color-luminosity diagrams have a thin line of stars running right up

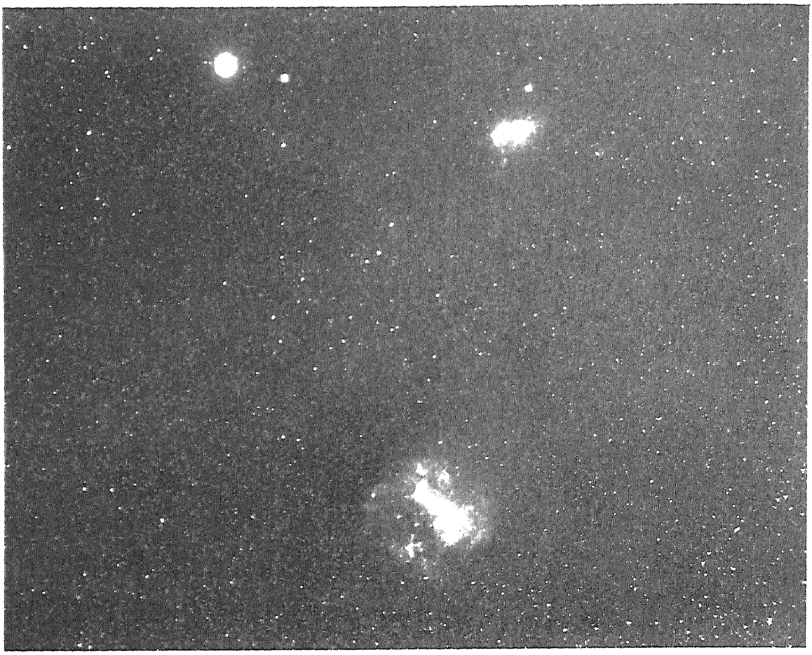


Figure 15-6 A photograph showing the Magellanic Clouds. Close to the Small Cloud is the globular cluster 47 Tucanae, which is a member of our galaxy. (Harvard College Observatory)

the main sequence to the blue giants. Other globular clusters in the Large Cloud appear to be a billion years old, but still younger than any in our galaxy. It would appear as if star formation in the Magellanic Clouds has either started late or for some reason been retarded.

The Clouds have radial velocities of recession of 170 and 100 miles per second respectively when referred to the sun, which moves within our own galaxy. When their motions are referred to the center of our galaxy, however, we find that the Clouds have essentially zero radial velocity. This fact, coupled with their nearness to our galaxy, has led some astronomers to believe that they are revolving about our galaxy as satellites. Radio observations of the 21-cm line of neutral hydrogen indicate that the two Clouds are enclosed in a common envelope of hydrogen gas. But unlike some systems (Figure 15-7), there does not appear to be a bridge of stars connecting the Magellanic Clouds to our galaxy. The Andromeda galaxy also has two satellite galaxies, but both of these are elliptical galaxies.

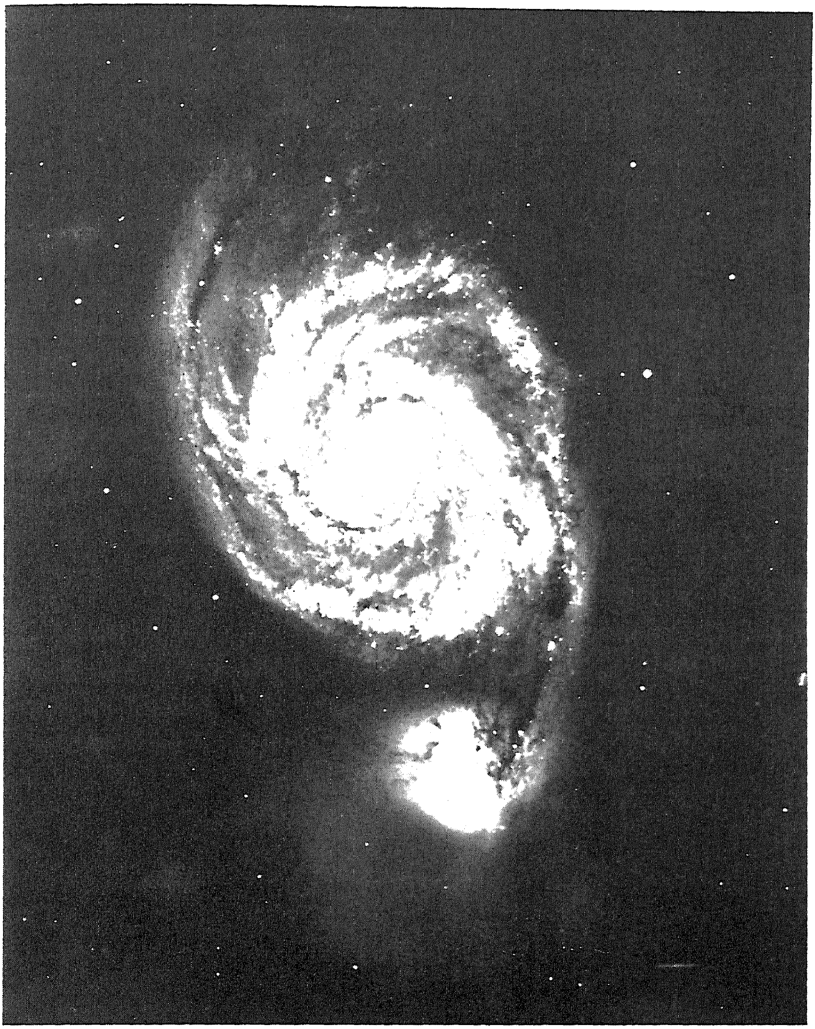


Figure 15-7 Spiral galaxy NGC 5194 with a bridge of stars connecting it to its satellite galaxy NGC 5195. (Mount Wilson and Palomar Observatories)

It has become clear that the most common galaxy of all is the *dwarf galaxy* (Figure 15-8). These galaxies are difficult to detect, for they are not as obvious nor as luminous as the bigger galaxies, yet they exceed the bigger ones in number. There are more dwarf galaxies in the immediate neighborhood of our galaxy than all other kinds put together. Large numbers of dwarf galaxies have been detected in nearby clusters of galaxies. Some of these dwarf galaxies

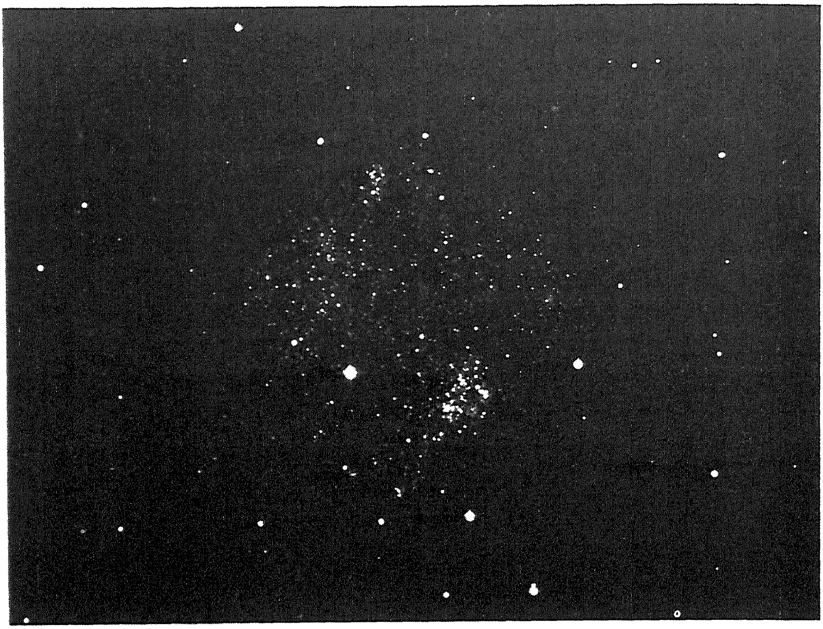


Figure 15-8 The dwarf galaxy in Sextans. (Mount Wilson and Palomar Observatories)

are regular in shape, and these tend to have Population II stars; the irregularly shaped dwarfs tend to have very luminous blue stars.

Very Active Galaxies

The discussion of galaxies so far has not introduced any new element into our concept of stellar systems. Population I objects appear in spiral arms, Population II objects in nuclei and elliptical galaxies. Galaxies rotate, but they move rather majestically. Even with an occasional supernova, galaxies seem to be rather placid; but this is not a safe conclusion. The galaxy M 82 (Figure 15-9*a*) was once classified as a peculiar galaxy; it certainly appears to be neither an elliptical nor a spiral galaxy. As optical identification of discrete radio sources proceeded, however, it became clear that M 82 is a very strong radio source. Subsequent observations have revealed that it is far from placid. A photograph taken with only the light of the red Hα line of hydrogen reveals a fine structure of filaments extending outward from the nucleus in a direction perpendicular to what appears to be the plane of the galaxy (Figure 15-9*b*). Since the plane of the galaxy is inclined by about 8° from the line of sight, observations

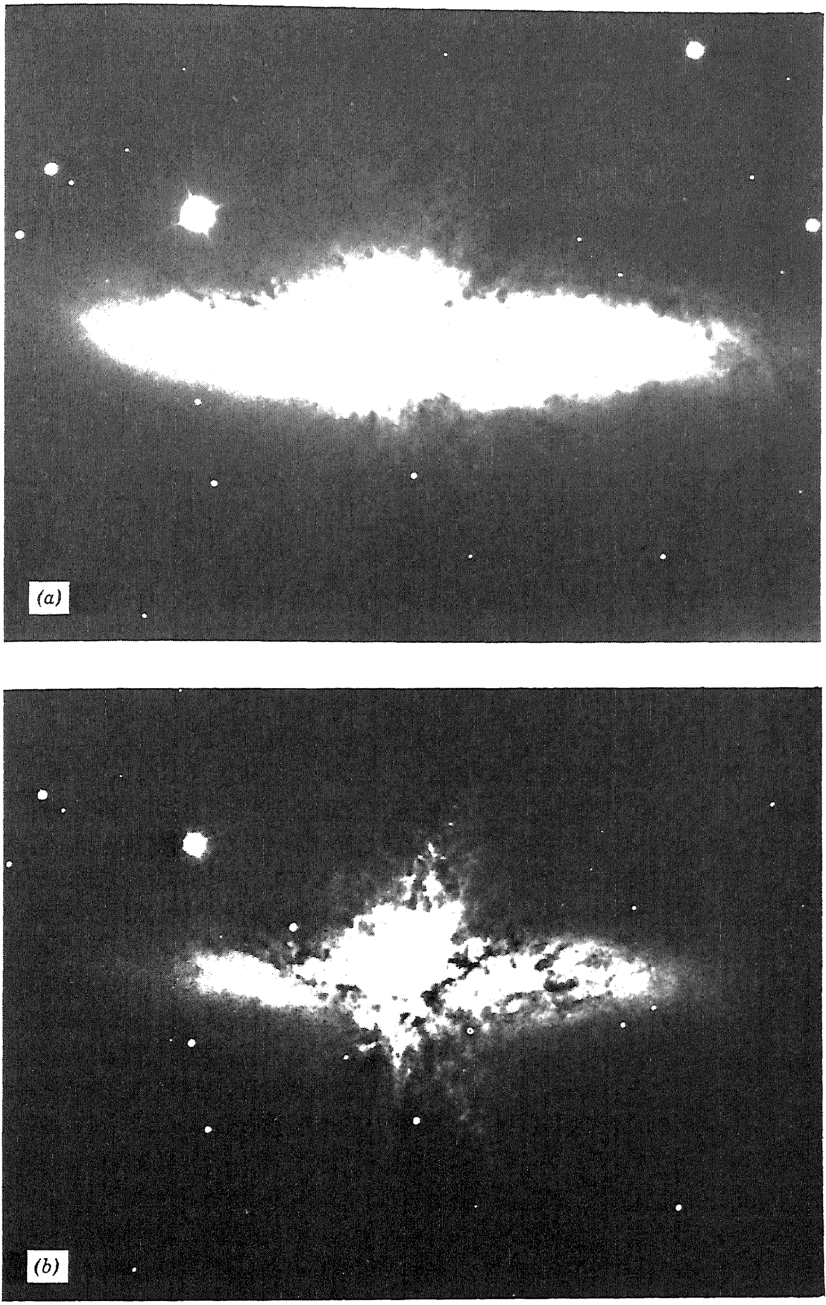


Figure 15-9 Exploding galaxy M 82 photographed in (*a*) white light, and (*b*) red light of hydrogen. (Mount Wilson and Palomar Observatories)

of radial velocity of these gases can be made. Their velocity indicates that the gases have been ejected by the nucleus. Furthermore, the velocity of the ejected gases increases in proportion to their distance from the center of the nucleus. The gases near the ends of the filaments are traveling outward at a velocity of 600 miles per second, and judging from their present distance—14,000 light years from the nucleus—they must have left the nucleus about 1.5 million years ago. The gases only half as far out have just one-half that velocity, so they, too, must have left the nucleus 1.5 million years ago. In fact, it seems that all of the gas left at the same time—the galaxy suffered an explosion!

If those gases have been slowed down by their interaction with intergalactic gases as well as by the galaxy's gravitational field, then the explosion may have occurred 2 or 3 million years ago. In any event it was only a short time ago on the astronomical scale. If, however, the inclination of the galaxy is not 8°, then the velocity of the gases is not that given, nor is the time since the explosion correct.

These ejected gases emit not only a bright-line spectrum but a continuous spectrum as well. The light of the continuous spectrum is polarized. This observation coupled with the fact that M 82 is a strong nonthermal radio source leaves little doubt that this galaxy emits synchrotron radiation. It would appear as if the magnetic field passes through the nucleus parallel to the axis of rotation of the galaxy, and the high-speed electrons are spiraling about the magnetic lines of forces as they proceed outward from the nucleus.

Another galaxy, M 87, when photographed on a long-time exposure appears as a normal elliptical galaxy (Figure 15-10*a*). But it, too, proves to be a source of strong radio energy, so it was investigated more thoroughly. A short-time exposure revealed a jet of gas extending outward from the nucleus (Figure 15-10*b*). This galaxy is also a source of X-ray radiation, indicating that the jet must contain very high-energy electrons indeed. The galaxy NGC 5128 (Figure 15-11*a*) is also a radio galaxy, but the radio energy does not all originate in the optical portion of the galaxy. A great deal of the radio energy comes from invisible gases on either side of and several diameters from the center of the galaxy (Figure 15-11*b*).

Details of these unusual galaxies are still being sought; the reasons why these galaxies emit such vast quantities of energy will become clear only after more details have been acquired, and discussed among the astronomers and then digested. Then perhaps a fresh idea will

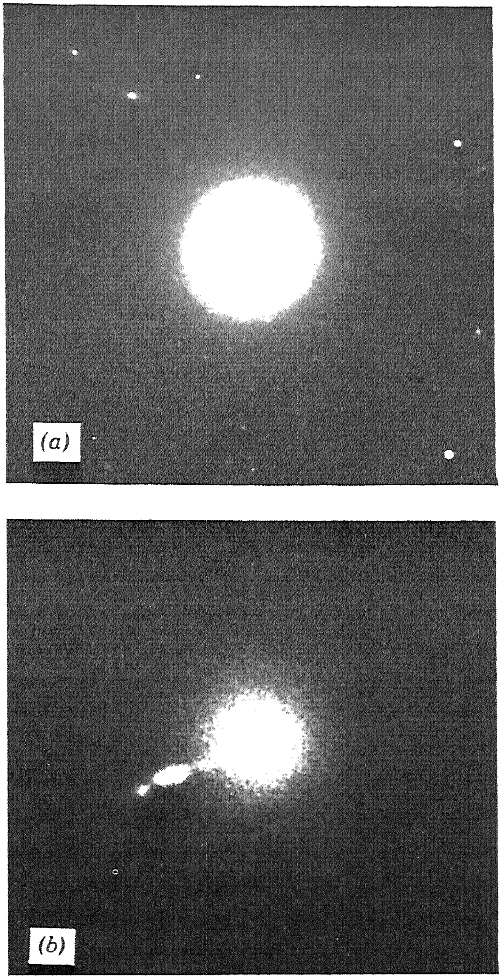


Figure 15-10 An elliptical galaxy, M 87 with (*a*) a long exposure, and (*b*) a short exposure revealing the jet-like source of radio energy. (Lick Observatory)

appear which may serve to describe these exploding (if they are) galaxies.

The Milky Way Galaxy

Which, if any, of these many and varied galaxies does our own Milky Way galaxy resemble?

Since our galaxy has a great deal of interstellar gas and dust and emission nebulae, it cannot be an elliptical galaxy. Since it has a very definite plane of symmetry, namely the galactic plane as defined

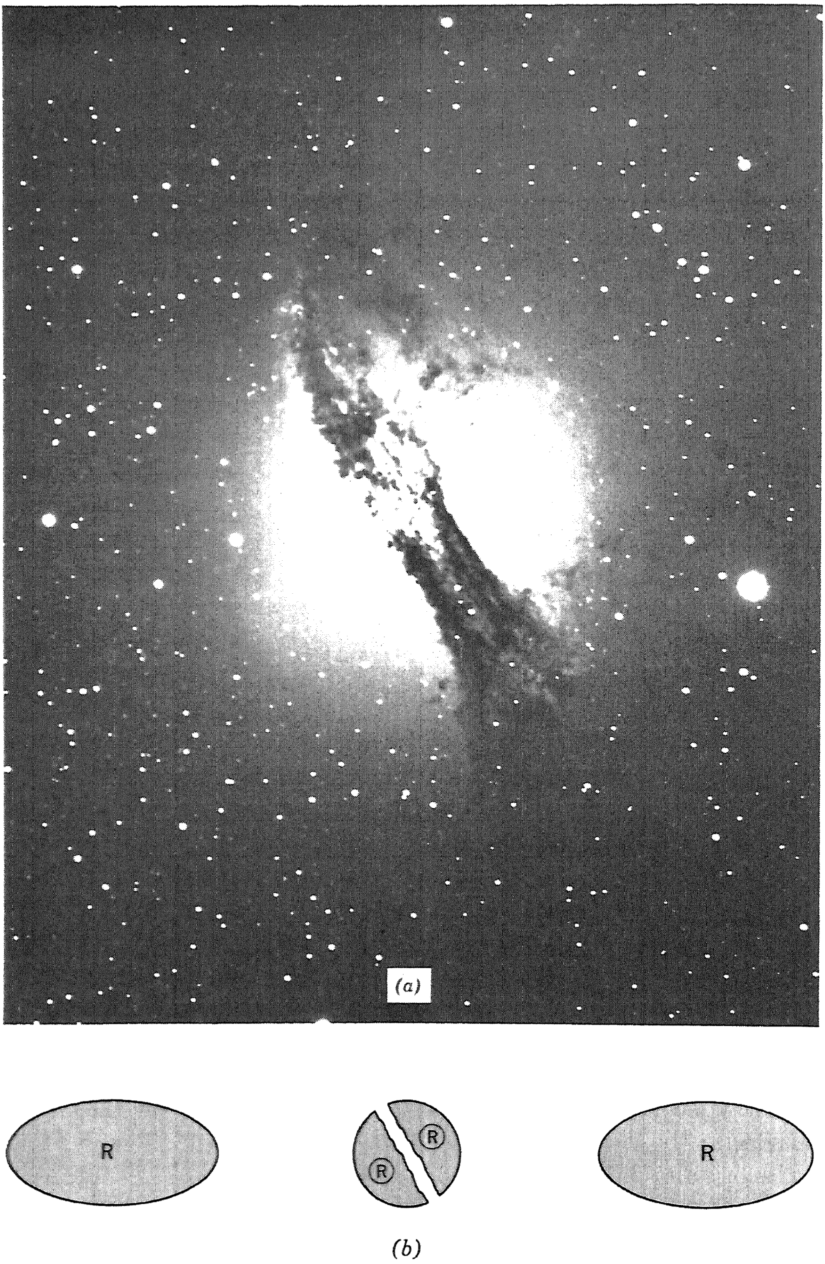


Figure 15-11 A strong radio galaxy NGC 5128 in Centauris. (*a*) photograph, (*b*) regions marked *R* indicate where the radio energy originates. (Mount Wilson and Palomar Observatories)

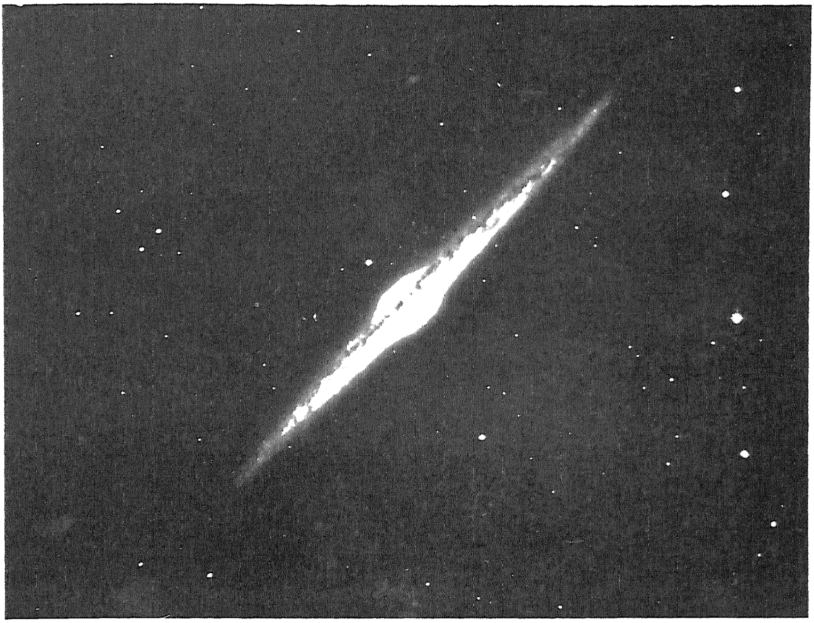


Figure 15-12 Spiral galaxy NGC 4565 seen "edge on." (Mount Wilson and Palomar Observatories)

by the Milky Way, it does not appear to be an irregular galaxy. We are left, therefore, with the tentative conclusion that our galaxy has a spiral structure. This conclusion is in harmony with a cursory glance at our galaxy, the shape of which can best be imagined by looking at the Milky Way, which forms a fairly narrow band around the celestial sphere. It must therefore be a fairly flat system, perhaps similar in cross section of the spiral galaxy in Figure 15-12 which is seen "edge on." The similarity is made more apparent by comparing this galaxy (NGC 4565) with a wide-angle photograph of the Milky Way (Figure 15-13). The difference between the two photographs, of course, is that one is viewed from the outside and the other is viewed from the inside.

Ours, then, is a spiral galaxy. We want to determine its size and structure—that is, locate the galactic nucleus (locate the position of the sun relative to the nucleus), observe evidence of spiral arms, and determine its mode of rotation.

In finding the nucleus of our galaxy we automatically locate the relative position of the sun. This has been done by several methods, one of which is based on the location of globular clusters (see p. 361).

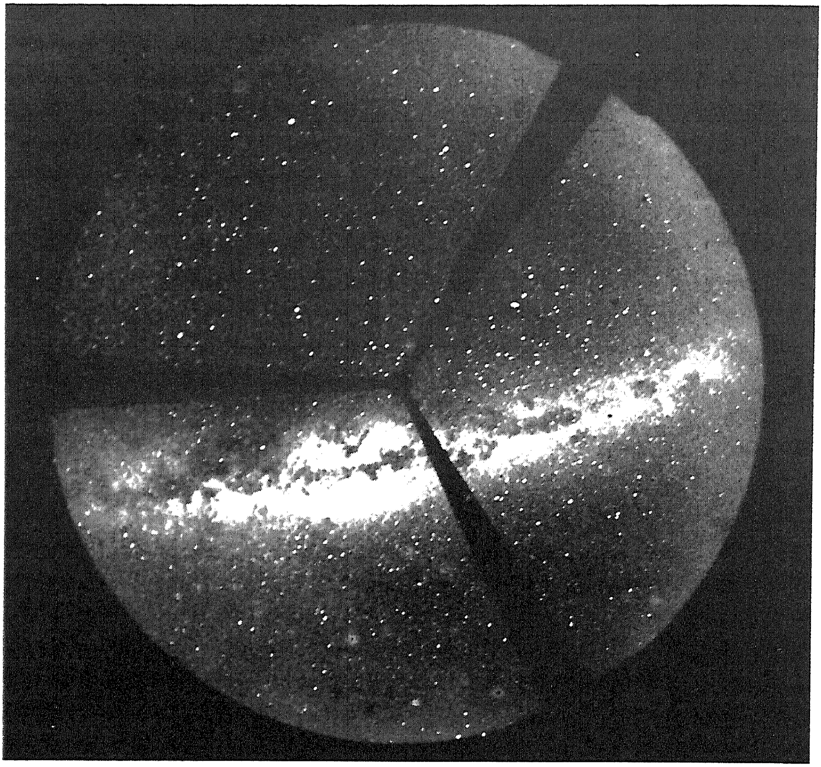


Figure 15-13 Wide-angle photograph of the Milky Way. (Washburn Observatory)

Of the 121 known globular clusters about 30 are concentrated in the constellation of Sagittarius. The remaining 90 are spread over a much larger part of the sky, the vast majority of them being in that half of the celestial sphere that is centered on Sagittarius. The distances of these globular clusters can be determined with a fair degree of accuracy, for not only do they contain RR Lyrae stars but many also inhabit regions of the sky that are relatively free from interstellar material.

Similar studies have been made of RR Lyrae stars not in globular clusters. These RR Lyrae stars form a system similar to that formed by the globular clusters but not as spherical; it is flattened slightly in the direction of the galactic plane. Other studies have been made of the location of Cepheids, of O and B stars; radio studies have been made to determine the distribution of interstellar hydrogen in our galaxy; and the best estimates of the distance of the sun from the center is 33,000 light years. The sun is about 40 light years from the galactic plane. (Our Earth is so oriented that its south pole is closer to the galactic plane than its north pole.) The studies have led to the further conclusion that our galaxy has a diameter of about 100,000 light years, which places the sun about two-thirds of the way out from the center. The region of the galaxy in which the sun resides is about 5,000 light years thick; thus the sun lies relatively close to the galactic plane. From studies of the size and rotation of our galaxy it is estimated to contain about 200 billion stars.

Having located the nucleus of our galaxy we now have to find evidence of spiral arms. Spiral arms in other galaxies are composed mostly of Population I objects: O and B stars, emission nebulae, interstellar clouds of neutral and ionized hydrogen, open clusters, and Population I Cepheids. There are other Population I objects such as T Tauri stars and flare stars, but these objects are relatively faint and therefore do not serve us well in locating distant spiral arms, although they are useful in verifying near-by spiral arms.

In order to locate spiral arms we look for regions in the galaxy where Population I objects have a rather high space density. This means that we will have to locate the objects in space, that is, give directions in terms of some reference system plus distances from the sun. In effect we will be locating objects on a three-dimensional graph.

The reference circle that is the most convenient for studies of the galaxy is the *galactic equator*—the intersection of the galactic plane with the celestial sphere. Locating the galactic equator, however, is not an easy task. A rough estimate of its location can be made by bisecting the Milky Way or it can be more precisely located by studying the distribution of certain objects over the celestial sphere. The objects that give the best results are the ones that can be seen for the greatest distance and therefore include the globular clusters, the O and B stars, the faint stars, and the Cepheids. The distribution of each of these is such that their number increases toward the galactic equator. Radio studies of the distribution of hydrogen gas, both ionized and neutral, have been invaluable in locating the central plane of our galaxy.

Having established the galactic equator that bisects the Milky Way we have the reference circle for a system of coordinates called *galactic latitude* and *galactic longitude*. Galactic latitude is measured in degrees north and south of the galactic equator and extends to the galactic poles. By definition the north galactic pole lies in that hemisphere which includes the north celestial pole, although the two poles are about 62° apart.

Galactic longitude must be measured from some chosen point on the galactic equator. At a meeting of the International Astronomical Union in Moscow in 1958, the point from which galactic longitude is measured was selected as that point on the celestial sphere which is defined as the galactic center.* The location of this point was determined by observations of the 21-cm line and lies in the direction of Sagittarius. Galactic longitude is measured from this point eastward along the galactic equator through 360°.

There are, therefore, three coordinate systems on the celestial sphere. These are based on the three principal planes: the plane of the Earth's equator, the plane of the Earth's orbit about the sun, and the plane of the galaxy. The angle formed by the intersection of the galactic equator with the celestial equator is about 62°. Thus the plane of the Earth's equator makes an angle of 62° with the plane of the galaxy. The two planes intersect at two points in diametrically opposite parts of the sky. One of the intersections is in the constellation of Aquila, just south and west of the star Altair. The other intersection is just east of Orion, between Orion and the star Procyon.

Since we have established the galactic equator as the most convenient reference circle for locating stars in our galaxy, we can now proceed with our search for evidence of spiral arms.

To locate the direction of any object in the galaxy we must give its position on the galactic coordinate system, that is, its galactic latitude and longitude. This position will define the direction of the object with respect to the sun. But to locate the object in space we must also determine its distance from the sun. That is, we must make the transition from the two-dimensional celestial sphere to three-dimensional space.

Let us consider first the O and B stars, since they are very luminous. The intrinsic luminosity of an O or B star can be determined by placing it on the H-R diagram, and its apparent brightness can be determined by means of photoelectric measurements. Its distance, however, cannot be determined directly by the inverse square law (as discussed on p. 259), for astronomers have accepted the fact that they can see these distant stars only through the scattering and absorbing interstellar material. This scattering and absorbing makes the stars appear fainter and thus more distant than they would be were there no interstellar material.

---

* Previous to this the origin of galactic longitude was in the constellation of Aquila at one of the two intersections of the galactic equator with the celestial equator.

In order to determine the true distance as accurately as possible, some estimate must be made of the amount of light that has been lost in transit between the star and the Earth. One method is to determine the color excess of the star. The color excess, it will be recalled, is a measure of the amount of light scattered by interstellar material. The greater the color excess the more light lost in transit, and the farther away the star appears. The amount of light absorbed can be estimated from the amount scattered. By considering both of these factors, astronomers are able to make reasonable estimates of stellar distances by applying corrections to the inverse square law.

When the distances and directions of a large number of distant O and B stars have been determined we can see that they are not distributed randomly about the sun but fall into patterns that strongly suggest the existence of spiral arms. This indication of spiral structure has been verified by studies involving the space distribution of emission nebulae, which when coupled with the work on O and B stars and Population I Cepheids discloses the spiral features shown in Figure 15-14.

According to Bart J. Bok of Mount Stromlo Observatory in Australia the sun resides on the inner edge of the Carina-Cygnus arm (a spur of which contains the Orion complex) which extends from the constellation of Cygnus through Cassiopeia, Perseus, Orion, Monoceros, to Carina in the southern hemisphere. This arm includes such notable features as the Coalsack, many O and B stars in associations, young galactic clusters, and emission nebulae. The Perseus arm includes the double cluster of Perseus, although it does not contain as many emission nebulae as does the Carina-Cygnus arm. The Sagittarius arm, which is closer to the galactic center than the Orion arm, contains an abundance of emission nebulae and the star clouds in Sagittarius (Figure 14-4). Additional evidence for the spiral structure of our galaxy is given later in this chapter.

Rotation of Our Galaxy

We have therefore reached the conclusion that our galaxy is spiral in structure and we can now investigate the problem of rotation. If the entire galaxy rotates like a solid wheel, none of the stars will move with respect to one another, and we should notice no radial velocities or proper motions among the stars. Similarly, two people

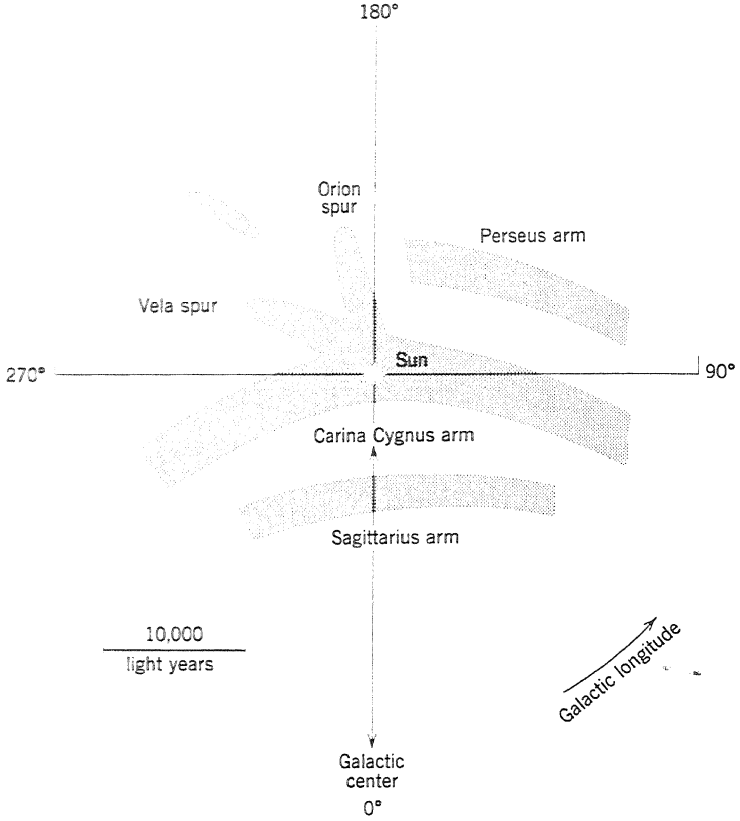


Figure 15-14 **Studies of the space distribution of O and B type stars and emission nebulae lead to the impression of spiral structure in the vicinity of the sun. (After Bart J. Bok)**

sitting on stationary horses on a merry-go-round do not move with respect to each other even though the merry-go-round is rotating. If, on the other hand, the stars revolve about the center of our galaxy with motions comparable to the motions of the planets about the sun, we should definitely observe not only radial velocities but also proper motions.

Figure 15-15*a* shows three circular orbits concentric with the galactic center. If the stars revolve in these orbits in a manner that satisfies Kepler's third law, the inner star will travel faster than the other two stars. The middle star (our sun) will have an intermediate orbital velocity, and the outer star will be the slowest. The orbital velocity of each is represented by the length of the arrow.

With the sun in the position shown we would be able to detect a Doppler shift to the violet (negative radial velocity) for an inner star at *a*, as it is overtaking the sun because of its higher orbital velocity. We would observe the same Doppler shift for an outer star at *f*, since the sun is overtaking it. On the other hand, a Doppler shift to the red (positive radial velocity) would be observed for an inner star in position *c*, since it is leaving the sun behind. A similar shift would be observed for an outer star at *d*, for the sun is leaving it behind.

If inner and outer stars are at *b* and *e* respectively, their spectra would exhibit no Doppler shift. The orbital motion of each of these two stars, however, would be recognizable by their proper motions. The inner star would seem to travel in the direction indicated by the arrow because it is passing the sun, while the outer star would appear to travel in a direction opposite that of the arrow because it is falling behind the sun. A star at *b* has a galactic longitude of 0°, but since it is moving faster than the sun it will eventually exhibit a Doppler shift.

Finally, if there were stars in the same orbit as the sun in either positions *g* or *h* they would exhibit no Doppler shift since they would be near the sun and traveling with the same speed. Thus, if the galaxy rotates according to Kepler's laws, a star's radial velocity will depend in part on its galactic longitude.

When the actual radial velocities of the distant O and B stars are examined, we find that in general they do indeed depend on the star's galactic longitude, although each star exhibits its own individual motion in addition to this over-all galactic rotation. The O and B stars at galactic longitude 45° (see Figure 15-15*a*), in general, exhibit a positive radial velocity indicating recession. At 135° galactic longitude they exhibit a negative radial velocity, at 225° a positive radial velocity, and at 315° a negative radial velocity. This evidence further substantiates the belief that our galaxy is a spiral, for the rotation of its outer regions is Keplerian (that is, it is similar to the planetary motion about the sun) and thus resembles the rotation of other spiral galaxies.

The amount of a star's radial velocity resulting from galactic rotation depends not only on its galactic longitude but also on its distance from the sun. Radial velocities increase for stars ever more distant from the sun until the position *T* (Figure 15-15*b*) is reached. At this point a given star's radial velocity is at a maximum, since it is moving directly away from the sun (and at position *T′*, directly

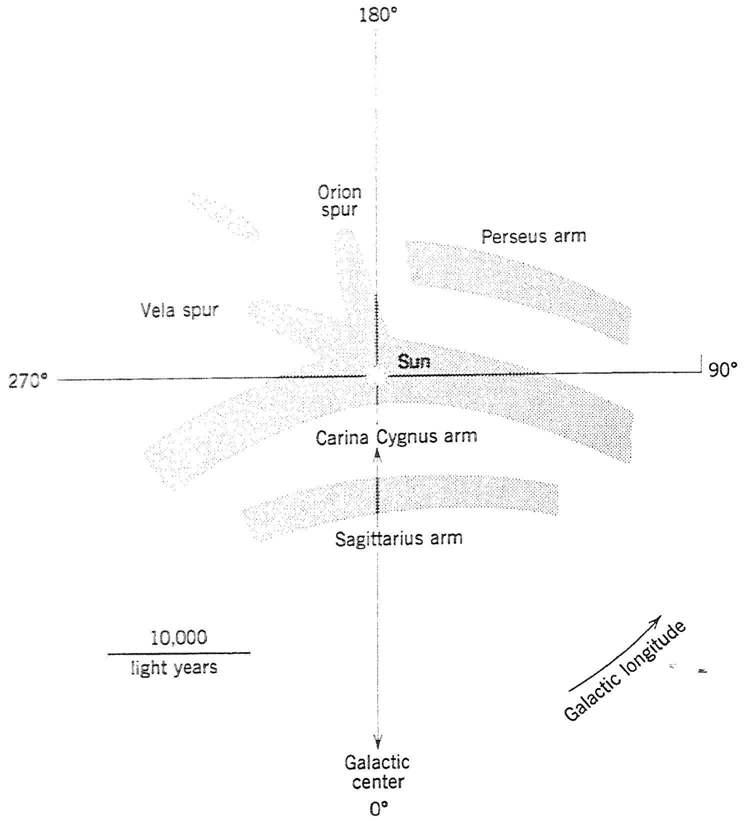


Figure 15-14 **Studies of the space distribution of O and B type stars and emission nebulae lead to the impression of spiral structure in the vicinity of the sun. (After Bart J. Bok)**

sitting on stationary horses on a merry-go-round do not move with respect to each other even though the merry-go-round is rotating. If, on the other hand, the stars revolve about the center of our galaxy with motions comparable to the motions of the planets about the sun, we should definitely observe not only radial velocities but also proper motions.

Figure 15-15*a* shows three circular orbits concentric with the galactic center. If the stars revolve in these orbits in a manner that satisfies Kepler's third law, the inner star will travel faster than the other two stars. The middle star (our sun) will have an intermediate orbital velocity, and the outer star will be the slowest. The orbital velocity of each is represented by the length of the arrow.

With the sun in the position shown we would be able to detect a Doppler shift to the violet (negative radial velocity) for an inner star at *a*, as it is overtaking the sun because of its higher orbital velocity. We would observe the same Doppler shift for an outer star at *f*, since the sun is overtaking it. On the other hand, a Doppler shift to the red (positive radial velocity) would be observed for an inner star in position *c*, since it is leaving the sun behind. A similar shift would be observed for an outer star at *d*, for the sun is leaving it behind.

If inner and outer stars are at *b* and *e* respectively, their spectra would exhibit no Doppler shift. The orbital motion of each of these two stars, however, would be recognizable by their proper motions. The inner star would seem to travel in the direction indicated by the arrow because it is passing the sun, while the outer star would appear to travel in a direction opposite that of the arrow because it is falling behind the sun. A star at *b* has a galactic longitude of 0°, but since it is moving faster than the sun it will eventually exhibit a Doppler shift.

Finally, if there were stars in the same orbit as the sun in either positions *g* or *h* they would exhibit no Doppler shift since they would be near the sun and traveling with the same speed. Thus, if the galaxy rotates according to Kepler's laws, a star's radial velocity will depend in part on its galactic longitude.

When the actual radial velocities of the distant O and B stars are examined, we find that in general they do indeed depend on the star's galactic longitude, although each star exhibits its own individual motion in addition to this over-all galactic rotation. The O and B stars at galactic longitude 45° (see Figure 15-15*a*), in general, exhibit a positive radial velocity indicating recession. At 135° galactic longitude they exhibit a negative radial velocity, at 225° a positive radial velocity, and at 315° a negative radial velocity. This evidence further substantiates the belief that our galaxy is a spiral, for the rotation of its outer regions is Keplerian (that is, it is similar to the planetary motion about the sun) and thus resembles the rotation of other spiral galaxies.

The amount of a star's radial velocity resulting from galactic rotation depends not only on its galactic longitude but also on its distance from the sun. Radial velocities increase for stars ever more distant from the sun until the position *T* (Figure 15-15*b*) is reached. At this point a given star's radial velocity is at a maximum, since it is moving directly away from the sun (and at position *T′*, directly

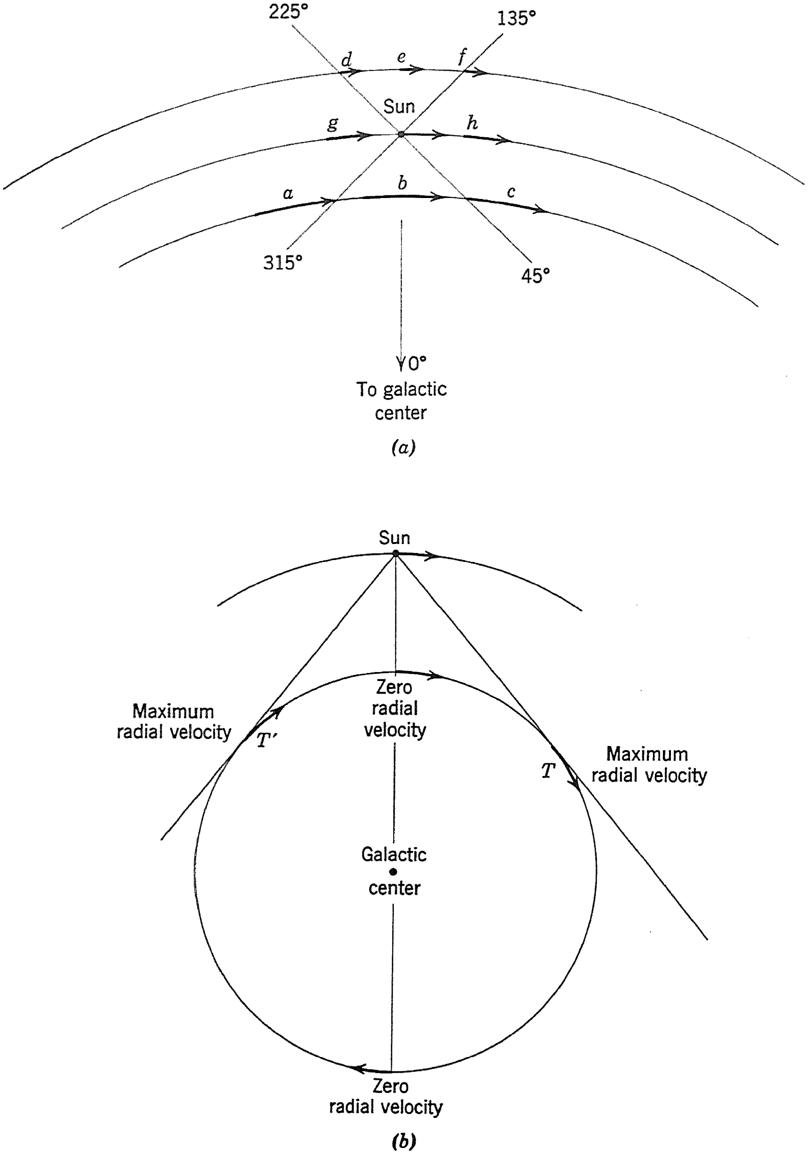


Figure 15-15 (*a*) The velocities of stars in the neighborhood of the sun showing Keplerian rotation. (*b*) The radial velocity of stars resulting from rotation of the galaxy reaches a maximum at points *T* and *T′*. If the stars travel in circular orbits their radial velocity would be zero at galactic longitude 0° and 180°.

toward the sun). Radial velocities should then begin to decrease until they reach zero for those stars in our galaxy that are diametrically opposite the sun.

We can see O and B stars along the galactic equator for a distance of only about 15,000 light years and since our sun is 33,000 light years from the center, the optically observed O and B stars do not include a very large portion of our galaxy. Radio observations, however, have extended this limit of observation to nearly the entire galaxy, and since radial velocity resulting from galactic rotation depends on the location of an object in our galaxy, radio astronomers find themselves in the unique position of being able to determine the spiral structure of most of our galaxy. Their observations all hinge on the fact that the 21-cm line, like every spectral line, is subject to the Doppler principle.

Figure 15-16*a* indicates why the radio astronomer has been so successful. If clouds of neutral hydrogen emitting radio radiation at 21 cm are at *a* and *b*, each will have a different radial velocity and thus display a different Doppler shift. A radio telescope pointed in the direction of *a* will record the 21-cm line as it is shown in Figure 15-16*b*. The larger hump on the curve results from radio emission of the nearer cloud, which displays a larger Doppler shift to longer wavelengths than the more distant cloud at *b*, because *a*'s motion is carrying it *directly* away from the Earth. In some cases more than two humps are seen on the 21-cm line, indicating that there are more than two hydrogen clouds in a given direction, each at different distances from the sun.

By employing this technique for other parts of the galaxy, radio astronomers have been able to locate the positions of clouds of neutral hydrogen in almost the entire galaxy. Since neutral hydrogen is found mostly in spiral arms, the cloud positions also locate the spiral arms of our galaxy. The results are shown in Figure 15-17. The conclusion drawn from the work of radio and optical astronomers is that our galaxy is a normal spiral, rather compactly wound, and that it thus resembles Andromeda (Figure 15-3) more closely than it does M 33 (Figure 15-4). The final word has not yet been said on the structure of our galaxy; there are still many uncertainties.

There are a few details concerning the rotation of our galaxy that need to be mentioned to emphasize the uncertainties of our knowledge about galactic rotation. The stars do not actually revolve about the galactic center in neat concentric circular orbits. Their orbits are slightly elliptical; thus the distance of each star from the galactic

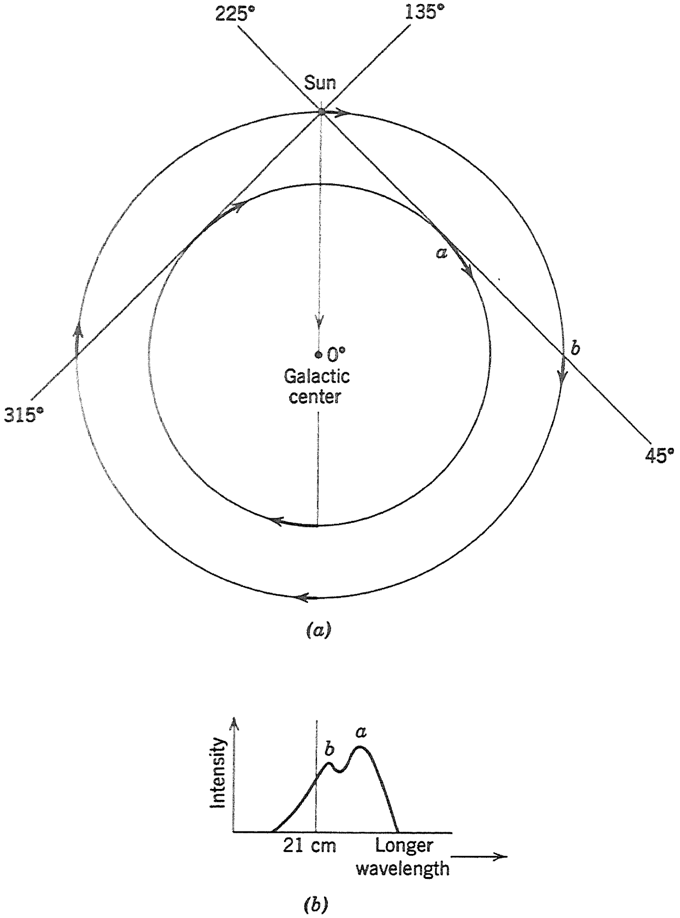


Figure 15-16 (*a*) The location of clouds of neutral hydrogen emitting the 21-cm line determines their radial velocity relative to the sun. (*b*) The 21-cm line may show two or more humps indicating two or more clouds of neutral hydrogen with different radial velocities.

center varies. If we observe stars at 0° and 180° galactic longitude, we do detect Doppler shifts, indicating motion away from and toward the galactic center. Thus each star has its own intrinsic motion superimposed upon the rotation of the galaxy. It is significant that this intrinsic motion is much larger for Population II stars than for Population I stars.

Recent studies of B stars and clouds of neutral hydrogen indicate that the central regions of our galaxy appear to be rotating much like a solid wheel, and the outer parts move more like the planets

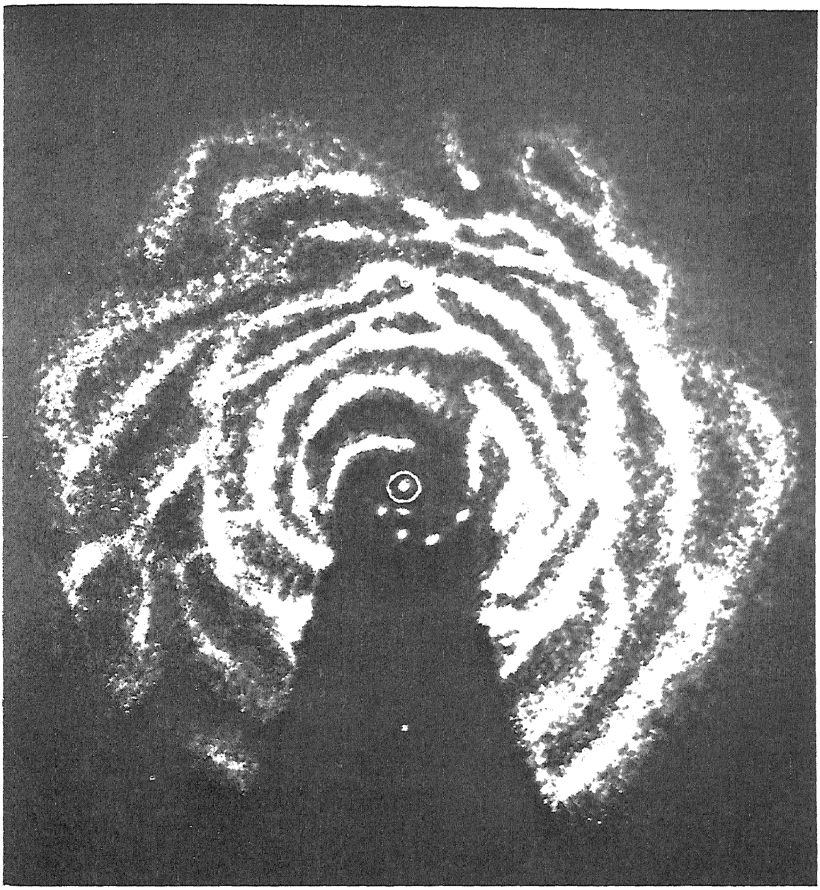


Figure 15-17 A drawing of the Milky Way galaxy as observed by the 21-cm line. The cross marks the center of the galactic nucleus, and the dot in the circle marks the position of the sun. This drawing was prepared by Westerhout of the Leiden Observatory, but the portion of the Milky Way visible only in the southern hemisphere was observed by Kerr in Australia. The portion omitted lies behind the galactic nucleus. (Leiden Observatory)

about the sun. This manner of galactic rotation, of course, is consistent with the rotation of the Andromeda galaxy.

Galactic rotation can be shown the most clearly if we plot the orbital velocity against the distance from the galactic center (Figure 15-18). If our entire galaxy were to rotate like a solid wheel, the orbital velocity of stars would increase in direct proportion with increasing distance from the center of the galaxy, as shown on the graph. If the rotation is Keplerian, the orbital velocities of the stars should decrease with increasing distance from the galactic center in accordance with the curve on the graph. The orbital velocity derived from the B stars is shown. In the region near the sun, this rotation is similar to "planetary" rotation, but it becomes more like "solid wheel" rotation closer to the center of the galaxy. The exact shape and slope of this curve is one of the uncertainties in galactic rotation.

It appears as if the hydrogen clouds, as observed by means of the 21-cm line, rotate in a manner different from the B stars. This seems to correspond to studies of star formation in nebula—either the stars acquire a velocity different from that of the nebula, or the velocity of the nebula changes after the stars have formed.

The fact that the apparent motion of a star (or group of stars) in the galaxy depends on its distance from the sun was used to make the first determination of luminosity for the Cepheids and RR Lyrae stars (see p. 278). Essentially, the method is to determine the average motion for a group of stars, let us say RR Lyrae stars, with respect to the sun. From this motion we can determine their average distance, and from their distance and their average brightness we can find their average luminosity. Since this method uses relative motions of groups of stars with respect to the sun, it has been called the method of *mean parallaxes.* Granted, it gives us only an average and would therefore be unworkable if many types of stars were considered together. It works, however, if stars of only one class are considered together; for example, we can determine the average luminosity of a group of A-type stars, or RR Lyrae stars, or Cepheids of a given period and Population.

The sun's motion about the center of the galaxy is of considerable interest to us, for to calculate the radial velocity of any object with respect to the galaxy we must subtract the sun's motion from the observed motion of the object. From studies of the motions of globular clusters, RR Lyrae stars, and near-by galaxies, it appears that the sun is moving around the center of the galaxy with an orbital velocity of about 200 miles per second. Its direction of motion is toward galactic longitude 90°, that is, at right angles to a line joining the sun with the center of the galaxy. The sun appears to travel in a circular orbit whose radius is 33,000 light years. At the observed velocity it will take the sun about 200 million years to complete one trip around its orbit. Most of the stars in the sun's neighborhood share this motion; the Population II stars, however, do not. Since most of the stars in the sun's neighborhood do share the sun's motion,

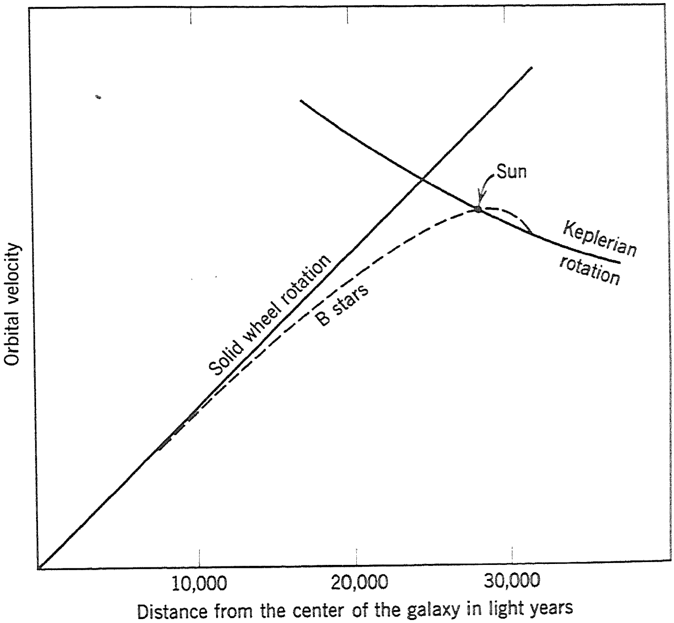


Figure 15-18 A graph plotting rotational velocity of our galaxy as a solid wheel, as Keplerian rotation, and by observations of B-type stars.

their velocity with respect to the sun is small—roughly about 20 miles per second. The population II stars, on the other hand, have velocities with respect to the sun of as much as 100 miles per second. This high relative velocity gave them the name *high-velocity stars,* a name which increased knowledge has made quite inappropriate.

The RR Lyrae stars are fairly representative of high-velocity stars and can be used to illustrate their motion. The average motion of the RR Lyrae stars is about 80 miles per second *with respect to the sun,* but in a direction opposite to the sun's motion. This means that they are moving about the galaxy more slowly than the sun is. If we subtract the sun's motion from their motion we find that they have orbital velocities of about 120 miles per second $(200 - 80 = 120)$.

Spiral Arms

One of the many unsolved problems of the structure of spiral galaxies is the very existence of the spiral arms. If galaxies rotate as they are observed to do, then each galaxy must have made many revolutions since its formation. Our sun, for example, is 4.6 billion

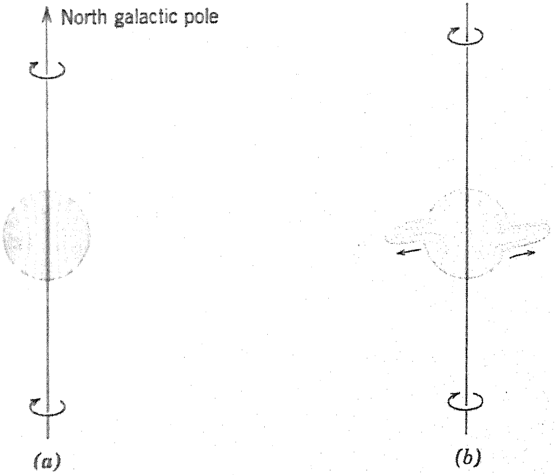
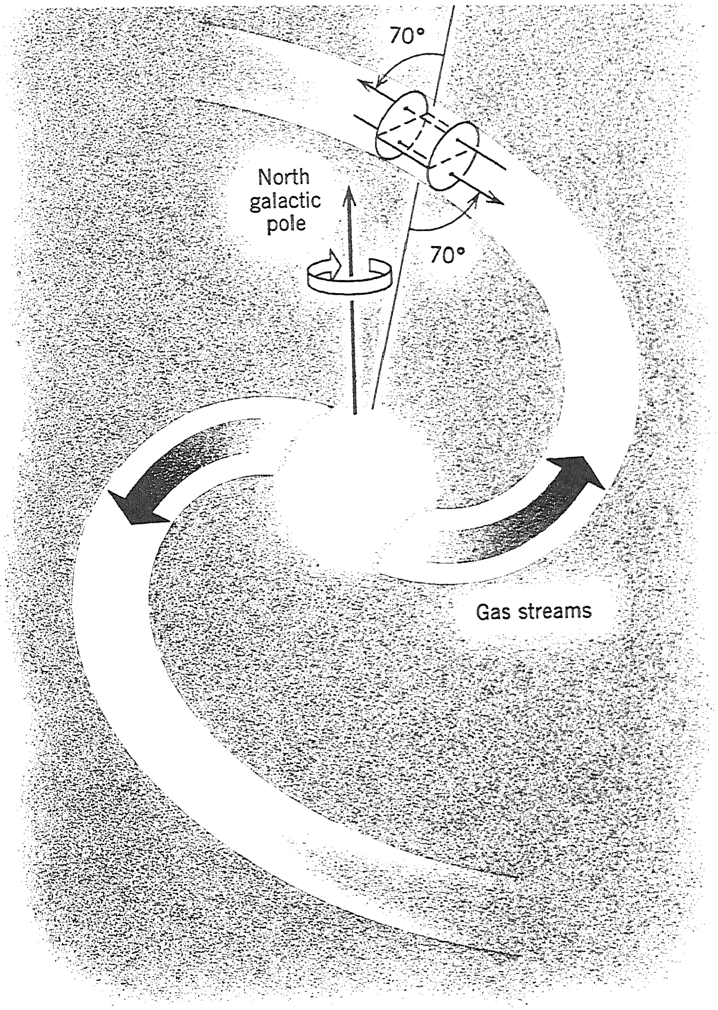


Figure 15-19 (*a*) The magnetic field of a very young rotating galaxy as proposed by Arp. (*b*) Perhaps as the result of an explosion, gas is ejected from the galaxy. The ionized gas carries the magnetic field with it. (*c*) The observed magnetic field in the spiral arm which contains the sun.

years old and it makes 1 revolution of the galaxy in 200 million years; it has, therefore, made a total of about 23 revolutions. The question now arises, why aren't the spiral arms twisted about the galaxy so tightly as to be unrecognizable? How do Sc galaxies maintain their open shape? What mechanism produces the stability which maintains the barred spiral arms? A different and yet related question concerns the existence of vast amounts of gas and very young stars in the spiral arms. Why hasn't all of this gas condensed into stars long ago? Why are stars still forming in a galaxy that is perhaps 10 to 20 billion years old?

Observations by radio telescopes of the 21-cm line indicate rather conclusively that hydrogen gas is being ejected by the nucleus of our galaxy (and the nuclei of other spirals as well). The amount of gas ejected is estimated to be about 1 solar mass per year, and it leaves the nucleus with a velocity of about 30 miles per second. Perhaps this is enough to replenish the gas which has formed the stars now visible, but by itself it does not account for the structure and stability of the spiral arms.

H. C. Arp of the Mount Wilson and Palomar Observatories has suggested that the spiral arms are maintained by the effect this ejected gas might have upon the magnetic field of the galaxy. That the galaxy



(c)

Figure 15-19 (Continued)

has an over-all magnetic field has been shown by a number of different observations. One of these observations is the polarization of light passing through dust particles which have been aligned by the galaxy's magnetic field. Presumably the magnetic field is oriented roughly as shown in Figure 15-19*a*.

A stream of ionized gas passing at right angles through a magnetic field will tend to carry that field right along with it. This has been observed in the interaction between the solar wind and the Earth's magnetic field (see p. 72). Arp proposed that the gases ejected by the nucleus carry the magnetic field of the galaxy right along with them forming a magnetic tube filled by the stream of ionized gas (Figure 15-19*b*). This magnetic tube could, it is proposed, give some strength to the spiral arms.

Actual observations of the magnetic field in the spiral arm in which the sun resides suggested and support this theory. In the northern portion of the arm the magnetic field points outward from the nucleus, in the southern portion it points toward the nucleus. This is precisely what Arp's theory specifies. Furthermore the magnetic field in the region of the sun points not directly to the nucleus but at an angle of 70° from it, verifying, if need be, the existence of an arm which does indeed spiral outward (Figure 15-19*c*).

Measurements of motions of stars and gases within a spiral arm indicate that the two classes of objects move differently. It would seem as if the stars form in the gases only to acquire a velocity different from the gas which nurtured them; the stars must leave the gases and populate the regions between the spiral arms. And between the spiral arms there has indeed been observed a substratum of red-dwarf stars. These stars must have left the spiral arms billions of years ago. The gas which forms these stars is then replenished by the ejection of gases from the nucleus. This leaves the question of how long can a nucleus supply gas to support spiral arms? Do galactic nuclei shrink with age? Do spiral arms then lose whatever strength they have?

Galactic Fossils

It is very striking indeed that there seems to be a significant relationship between the orbital motions and locations of stars, and their ages:

1. The globular clusters are the oldest members of our galaxy and they form a spherical system concentric with the galactic nucleus. Their orbital velocities are low, and their orbits are eccentric and inclined at large angles from the galactic plane.

2. The O and B stars are the youngest members and they form the flattest system, being nearly coincident with the galactic plane. Their orbits are very nearly circular and their orbital velocities are higher than those of any other stars.

3. The other stars seem to form a continuum between these two

extremes. At one extreme are the O and B stars, which are very young Population I stars. The late main-sequence stars (K and M) are usually older than the O and B stars. They form a system that is fairly flat, but their orbits are more likely to carry them farther from the galactic plane than do the orbits of the O and B stars. Their orbits are also less circular than those of the youngest stars, and their orbital velocities a little lower. The RR Lyrae stars are Population II stars but are younger than the globular clusters. They form a system that, although it extends far from the galactic plane, is slightly flattened and thus does not extend as far as the globular clusters. Their orbits are also eccentric and inclined from the galactic plane but not as much as the orbit of the globular clusters. Their orbital velocities are low but not as low as those of globular clusters.

The ages of stars are also related to their location in the galaxy in that the older stars are more likely to congregate near the galactic nucleus. Thus their orbits are more likely to be smaller than those of younger stars.

All these facts assembled are wonderful material for speculative thought; it appears that our galaxy, like a tree, has left rings that tell of its past. When our galaxy was at the age during which its globular clusters were forming from gaseous clouds of hydrogen, it must have had a shape quite similar to that now defined by the system of globular clusters—a nearly spherical shape. Furthermore, it must have rotated with a velocity nearly equal to that of the globular clusters—a low velocity. We speculate further that apparently the galaxy then began to contract, but the globular clusters, having formed, maintained their orbits. As the galaxy contracted it rotated more rapidly to conserve its angular momentum. In rotating more rapidly it began to flatten out and after roughly 10 billion years it has become the flat system that we now observe. The youngest stars and the remaining clouds of hydrogen are all part of this flat system; the older stars are not. With increasing age, the older stars form ever more spherical systems and on the average have smaller orbits and lower orbital velocities. Their orbits, however, can be more eccentric and more inclined from the galactic plane than those of younger stars. Three stages of galactic evolution are depicted in Figure 15-20.

This is not meant to imply that there are no Population II objects in the galactic plane, for there are. Some may even have orbits that lie in the galactic plane. Those Population II objects in orbits inclined from the galactic plane must still cross it twice in one revolution.

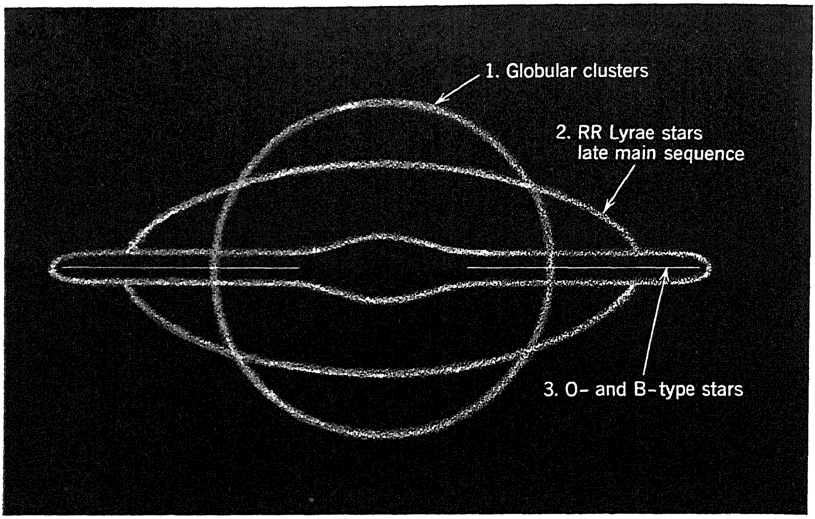


Figure 15-20 Speculation leads to these three stages of galactic evolution, because a star's location in our galaxy as well as its orbital inclination from the galactic plane and its speed of revolution about the galactic center all depend to a certain degree on the age of the star.

During this crossing they are temporarily (thousands of years) close to the galactic plane.

That so many globular clusters are concentrated near the galactic nucleus indicates that during the time when the globular clusters were forming, the shape of our galaxy may have resembled that of a small elliptical galaxy of type E0, with a high concentration of stars near the center, that gradually thins out with increasing distance from the center.

This comparison of our galaxy at this stage of development with elliptical galaxies is interesting; unfortunately, it is misleading. Although our galaxy in the early stages of its development may have had a shape similar to that of elliptical galaxies, it must have been different because elliptical galaxies contain little apparent interstellar gas and dust. Thus star formation must be nearly at an end in these Population II objects.

The terms Population I and Population II now take on added significance. They represent the extremes in age. There are stars of intermediate age and thus of intermediate Population class. Since we assume that star formation has been going on continually, the gradation from Population I to Population II must be continuous. The sun appears to be an old Population I star.

Evolution of Galaxies

Galaxies have a very strong tendency to form in clumps and clusters. Very often galaxies will form together in a very tight assemblage (see Figure 16-3). It is presumed that galaxies so close together are under the gravitational influence of one another. If this is true, then the relative masses of each galaxy can be estimated from their motions; the more massive galaxies will move the slowest, the least massive will move the fastest. So reasoned Thornton Page of Van Vleck Observatory, Wesleyan University, who then set about to determine the relative masses of many galaxies from a number of such close assemblages. Surprisingly enough he found that the E0 galaxies, the very spherical ones, are consistently nearly 30 times as massive as the open Sc spirals. Furthermore he found a gradation in mass from the E0 galaxies to the Sc galaxies. The dwarf galaxies, however, are not part of this gradation, their masses are even less than those of the Sc galaxies.

It is well known that elliptical galaxies do not rotate as fast as the spiral galaxies. However, the measure of rotation should include the amount of matter rotating, so we employ angular momentum as this measure. As in the formation of the solar system, the angular momentum of a galaxy should remain constant since no appreciable outside forces act to reduce it. But even with its greater mass the angular momentum of an elliptical galaxy is less than that of a spiral galaxy.

H. C. Arp plotted the mass of the galaxies against their angular momentum per unit mass and found a very interesting relationship (Figure 15-21). It is significant that the most massive galaxies have the least amount of angular momentum per unit mass; and that there should be a relatively uniform gradation between the E0 and the Sc galaxies. Arp suggests that this gradation results from the conditions present during the formation of the galaxies.

If the galaxies form from blobs of gas of nearly equal mass, then all galaxies should have equal mass; but if the gas from which each galaxy forms has a different amount of angular momentum, then the resulting galaxies might be different. Arp suggests that during the contraction of a large nonrotating blob of gas the density of the gas remained fairly uniform so that star formation proceeded at a fairly uniform rate and for only a relatively short period of time. All of the gas would have condensed into stars and all of the stars would have nearly the same age. In other words, an elliptical galaxy would form from a nonrotating blob of gas. Those blue-giant stars

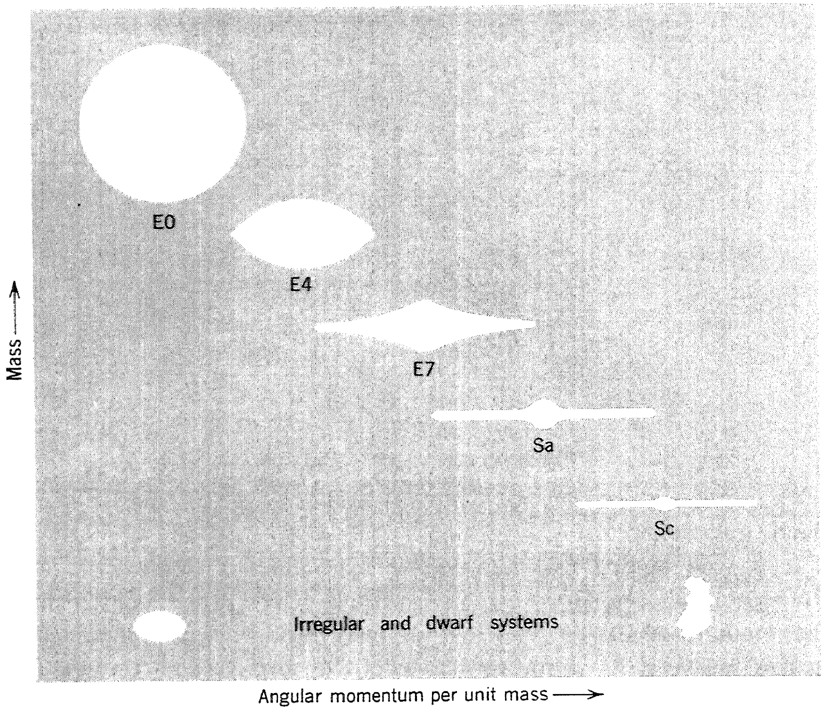


Figure 15-21 As the angular momentum per unit mass of galaxies increases, their total mass decreases.

formed during this early age would long since have evolved into white dwarfs. Their being no more gas to form new stars, no successive generations of blue giants would form.

On the other hand, if the initial blob of gas were rotating—if it had an appreciable amount of angular momentum—then as the gas contracted it would have rotated still faster to conserve angular momentum. As it rotated faster, gases would have been thrown out to form an equatorial bulge, and later a nebular disk reminiscent of the nebular disk about the proto-sun. The formation of the nebular disk would have decreased the density of the gas and consequently decreased the rate of star production. With the increased rotation much of the gas may have been thrown free of the forming galaxy resulting in a decrease in mass. Evidence for this idea is supported in the stages of evolution of our own galaxy.

Consequently, it is suggested that the amount of angular momentum of the initial blob of gas is the main factor determining whether the gas will become an E0, E1, E2, . . . , Sa, Sb, or Sc galaxy.

It is generally accepted that during the life-time of a spiral galaxy the production of supernovae and other gas-ejecting stars replenishes some of the interstellar material used up in star formation. If that initial interstellar material is nearly pure hydrogen, however, these gas-ejecting stars "contaminate" the interstellar gases. During the lifetime of a star hydrogen is converted into helium, which is in turn converted into carbon, oxygen, and the heavier elements. By ejecting gases back into space stars spew forth some of these heavier elements as well. Consequently, the first stars to form in a spiral galaxy should have a low percentage of metals; stars formed later in the life of a galaxy should have a higher percentage of metals. The most recently formed stars should have the highest percentage of metals of all the stars. It is well known that Population I stars have a higher percentage of metals than do the Population II stars, so the idea of an increasing metallic content of the interstellar gases is well supported. Presumably, all of the stars of the elliptical galaxies have the same, or nearly the same, chemical composition.

This suggestion of galactic evolution is one of the first to really bring some of the various shapes and forms of galaxies into a unified principle of formation. There are certainly many galaxies which do not fall into the straight sequence E0 to Sc, such as the barred spirals, the dwarf systems, the irregular systems, the exploding galaxies, etc. But more progress is made when astronomers have a framework—an hypothesis—upon which to work. If that hypothesis is later shown to be incorrect, astronomers have still learned and benefited from it. This suggestion of galactic evolution will serve astronomers very well, even if it cannot answer all of their questions now. The details will be forthcoming with additional studies, both observational and theoretical. This is the manner in which ideas develop into good descriptions of our universe.

Basic Vocabulary for Subsequent Reading

Barred spiral galaxy
Dwarf galaxy
Elliptical galaxy
Galactic equator
Galactic latitude
Galactic longitude
Irregular galaxy
Normal spiral galaxy
Population I objects
Population II objects

For Further Reading

Baade, Walter, *Evolution of Stars and Galaxies,* Harvard University Press, Cambridge, Mass., 1963.

Bok, Bart J., and Priscilla, *The Milky Way,* Harvard University Press, Cambridge, Mass., 1957.

Page, Thornton, ed., *Stars and Galaxies,* Prentice-Hall paperback, Englewood Cliffs, N.J., 1962.

Struve, Otto, and V. Zebergs, *Astronomy in the 20th Century,* chapters XIX, XX, Crowell Collier and Macmillan, New York, 1962.

Arp, H. C., "The Evolution of Galaxies," *Scientific American,* p. 71, January 1963.

Berge, G. L., and G. A. Seielstad, "The Magnetic Field of the Galaxy," *Scientific American,* p. 46, June 1965.

Bok, Bart J., "An International Discussion of Milky Way Research," *Sky and Telescope,* p. 4, July 1963.

Bok, Bart J., "The Large Cloud of Magellan," *Scientific American,* p. 32, January 1964.

Faulkner, D. J., "Current Studies of the Magellanic Clouds," *Sky and Telescope,* p. 69, August 1963.

Heeschen, D. S., "Radio Galaxies," *Scientific American,* p. 41, March 1962.

Hodge, Paul W., "Dwarf Galaxies," *Scientific American,* p. 78, May 1964.

Hodge, Paul W., "Globular Clusters in the Magellanic Clouds," *Sky and Telescope,* p. 72, February 1961.

Hodge, Paul W., "The Sculptor and Fornax Dwarf Galaxies," *Sky and Telescope,* p. 336, December 1964.

Page, T. L., "The Evolution of Galaxies," *Sky and Telescope,* part I, p. 4, January 1965; part II, p. 81, February 1965.

Struve, Otto, "Interacting Galaxies," *Sky and Telescope,* p. 16, January 1962.

Roberts, Morton S., "Hydrogen in Galaxies," *Scientific American,* p. 94, July 1963.

"Age of Our Galaxy," *Sky and Telescope,* p. 287, May 1964.

"Compact Galaxies," *Sky and Telescope,* p. 131, September 1964.

"The Heart of the Milky Way," *Sky and Telescope,* p. 282, May 1965.

"Helium and the Galaxy's Age," *Sky and Telescope,* p. 219, April 1965.

"Motions in the Galaxy M 77," *Sky and Telescope,* p. 143, March 1966.

"Nests of Supernovae," *Sky and Telescope,* p. 283, May 1965.

"New Advances: Sun, Space Astronomy, the Galaxy," *Sky and Telescope,* p. 264, November 1964.

"X-ray Galaxies," *Scientific American,* p. 50, April 1966.

A group of five galaxies with clouds of material connecting them. (Mount Wilson and Palomar Observatories)

Chapter 16 *The Universe and Relativity*

We define the universe as that which includes all matter in existence. Since the amount of matter and the size of the universe are not yet known, we must start with what we know about the *observable* universe. From their analyses of the observable universe astronomers are able to formulate hypotheses about the structure of the entire universe (the study of *cosmology*) and its origin, evolution, and even its future (the study of *cosmogony*).

The Distribution of Galaxies

The basic building block in the universe is the galaxy, just as the basic building block in a galaxy is the star. Astronomers observe all the visible galaxies projected onto the celestial sphere, and from this projection conclude that galaxies are not distributed at random throughout the universe.

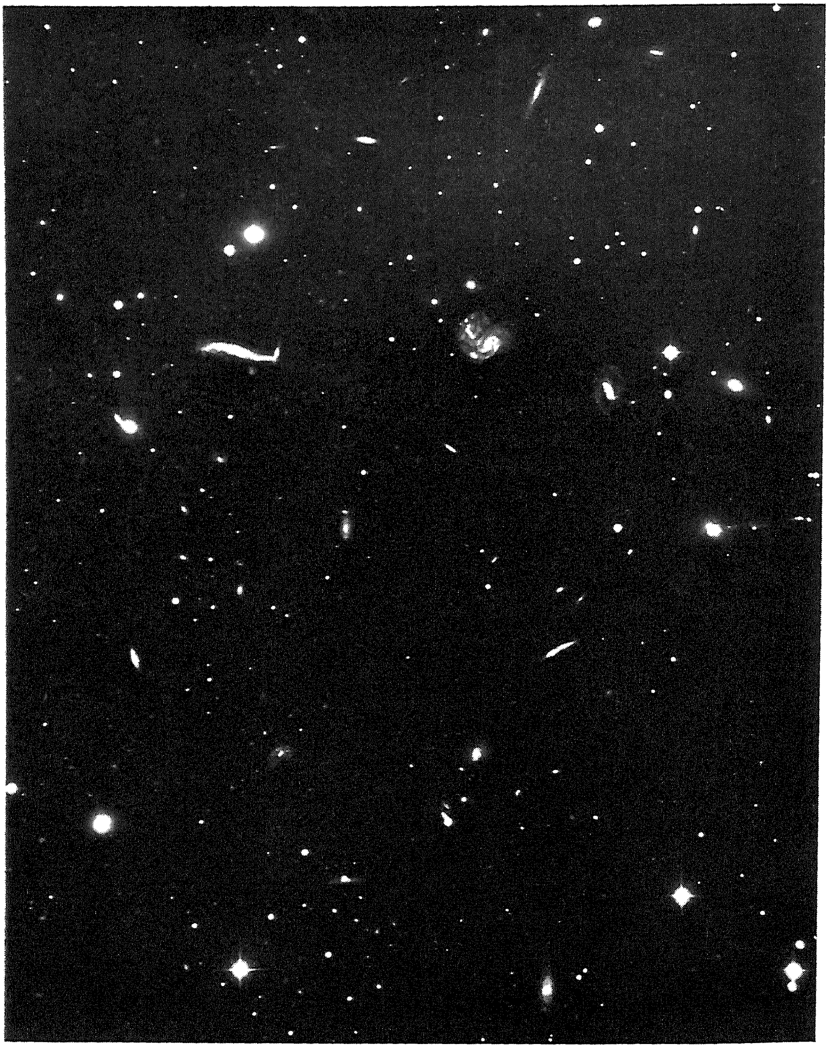


Figure 16-1*a* A cluster of galaxies in the constellation Hercules. (Mount Wilson and Palomar Observatories)

C. D. Shane, of the Lick Observatory, has presented conclusive evidence than the galaxies are clumped together into clusters, Figure 16-1*a*. The survey plates for the proper motion study at the Lick Observatory (see p. 245) contain thousands of galaxies as well as the foreground stars in our own galaxy. Shane and his co-workers counted the number of galaxies in each square degree of the sky

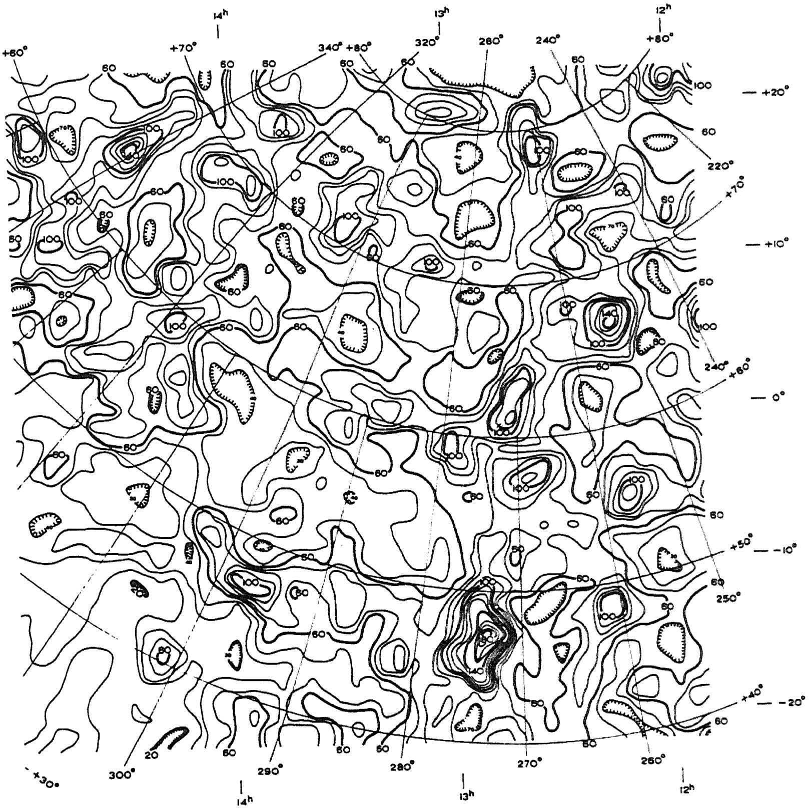


Figure 16-1*b* The distribution of galaxies in a portion of the celestial sphere as drawn by C. D. Shane indicates that galaxies tend to form in clusters. (Lick Observatory)

and plotted these counts on a large "celestial" map. They then drew contour lines joining those regions of the sky that had the same number of galaxies. A small portion of their "galactic contour map" is shown in Figure 16-1*b*. Each of the areas of closed contour lines represents a cluster of galaxies, and this is only a small portion of the entire celestial sphere! Although there also appear to be stray galaxies that are not contained in any cluster, they are the exception and not the rule.

Fritz Zwicky, of the Mount Wilson and Palomar Observatories, has concluded that many of these clusters have spherical symmetry; that the space density of the galaxies increases toward the center of the cluster; and that the very bright elliptical galaxies are concentrated toward the center, with the spiral and irregular galaxies more

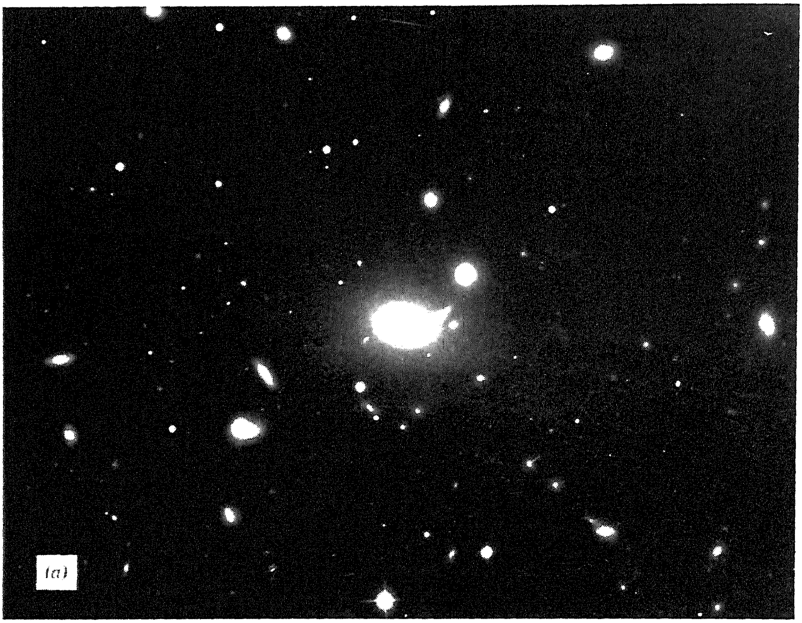


Figure 16-2 Photographs of clusters of galaxies all enlarged the same amount to show the decrease in apparent size and brightness of the galaxies with increasing distance. (*a*) Part of the Coma Berenices cluster estimated to be 136 million light years away. (*b*) Corona Borealis cluster estimated to be 400 million light years away. (*c*) Hydra cluster 1,100 million light years away. (Mount Wilson and Palomar Observatories)

prevalent in the outer regions. Some of the clusters may contain as many as 100,000 galaxies but these are exceptionally rich.

Our own galaxy is one of about 17 (there is some doubt of membership for a few) forming a rather small cluster called the *Local Group.* Our Local Group also contains the Magellanic Clouds, the Andromeda galaxy with its two satellites, M 33 (the spiral in Triangulum, Figure 15-4), and others. If the Large Magellanic Cloud is classified as a barred spiral, the Local Group contains all types of galaxies. The Andromeda galaxy and our galaxy are by far the largest members of this group; M 33 is a poor third, and the Magellanic Clouds are fourth and fifth in size. All these galaxies are included in a volume of space about 3 million light years in diameter.

Other well-known clusters of galaxies include: the Coma Berenices cluster with about 9000 members; the Corona Borealis cluster with at least 400 members; the Hydra cluster with over 200 members (Figure 16-2).

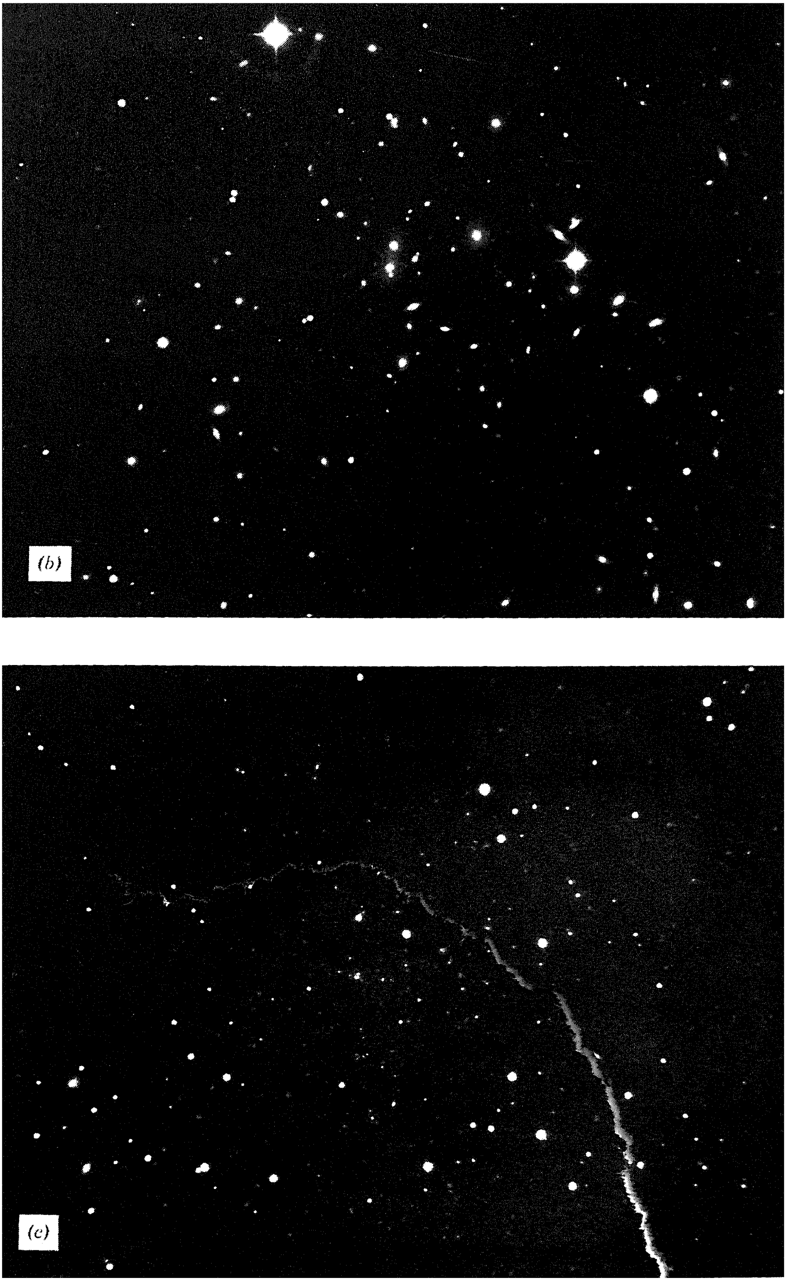


Figure 16-2 (Continued)

Each cluster of galaxies is a group bound within itself, held together by the mutual gravitational attraction of all the members. But this gravitational field does not restrict the motion of the member galaxies within the cluster; galaxies will have intrinsic velocities within the cluster of as much as 600 miles per second.

Zwicky has commented that each cluster appears to be an isolated unit and that no clusters form in pairs, groups, or clusters of clusters, which fact may have broad implications for Newton's law of gravitation applied to such large distances. Zwicky suggests that because of the large distances separating clusters of galaxies attraction ceases and repulsion sets in, for the clusters of galaxies seem to avoid one another.

This certainly cannot be said of the galaxies inside the clusters, for there is overwhelming evidence that galaxies tend to associate closely with one another. There are many galaxies that exist as pairs, others as three and more (Figure 16-3). The distance between galaxies in a cluster is not as great in comparison to their size as is the distance between the stars in comparison to their size. On the average, galaxies have a diameter of about 5×10^4 light years and in a cluster they have an average separation of roughly 30×10^4 light years. Thus they are separated, on the average, by a distance roughly six times their diameters, whereas stars in the region of the sun are separated by a distance of about 30 million solar diameters.

Zwicky amply shows that intergalactic gas and dust does exist. He points out that the number of faint clusters of galaxies is much reduced in regions of the sky covered by near-by rich clusters of galaxies. For example, there is an average of only 8 faint clusters seen through the central region of the Coma Berenices cluster in an area of the celestial sphere of 36 square degrees, but in regions where no near-by rich clusters exist there is an average of 70 clusters for the same area.

Distances of Galaxies

Although we can learn a great deal from the study of the distribution of galaxies on the two-dimensional celestial sphere, we must consider the universe as being three-dimensional and study the distances of the galaxies.

Let us review the ever-present problem of distance determination. For only the nearest stars does the method of heliocentric parallax do us any good. Beyond about 50 light years, heliocentric parallax

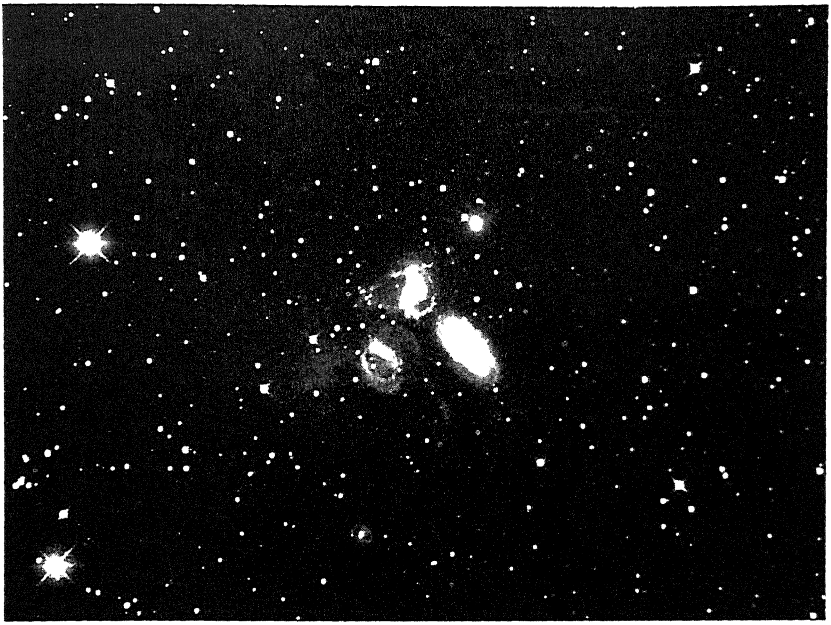


Figure 16-3 A group of five galaxies, called Stephan's Quintet, forming a compact system. (Mount Wilson and Palomar Observatories)

begins to fail and other less direct methods must be used. An astronomer can measure the apparent brightness of a star whose intrinsic luminosity is known, such as a Cepheid or RR Lyrae star. From the brightness and luminosity the distance can be determined by the inverse square law. The astronomer may have to account for the loss of light from scattering by interstellar matter, in which case his results are likely to be less accurate.

The distances of the nearest galaxies, the Magellanic Clouds, have been determined by using RR Lyrae stars, for their period-luminosity relation is more reliable than that of the Cepheids. The Cepheids must be used for more distant galaxies such as the Andromeda galaxy and M 33, since RR Lyrae stars cannot be seen in them. For still more distant galaxies not even the Cepheids can be seen and the astronomer must resort to other methods.

One method that has been used in the past is to estimate the luminosity of the brightest stars in a given galaxy. This method is based on the assumption that the brightest stars in each galaxy are of nearly the same luminosity. This luminosity can be estimated from studies of the stars in our galaxy as well as of the brightest stars in near-by galaxies of known distance. From this luminosity and the measured brightness of the stars the distance of the galaxy can then be estimated. But even this method fails with galaxies at still larger distances, so we must resort to other procedures.

Many galaxies are surrounded by a halo of globular clusters similar to ours which, because they contain thousands of stars, can be seen at greater distances than can the brightest stars in any galaxy. The brightest globular clusters can be used to estimate distances by the same method used for the brightest stars. Still more distant galaxies, however, are too far away for even the globular clusters to be seen, much less individual stars, except for an occasional supernova.

But an occasional supernova is of great help to the astronomer. As indicated on p. 295 supernovae can be used as a tool to estimate distance, since supernovae of either type I or II have fairly consistent luminosities within these types. Supernovae of type II reach luminosities of about 200 million suns and may thus be used to determine the distances of galaxies too far away for globular clusters to appear (Figure 16-4). Although supernovae, when observed, prove very valuable for the measurement of distances, they are sporadic and thus not readily available. Therefore other distance indicators are more frequently used.

The realization that galaxies form in clusters has led astronomers to try to use the galaxies themselves as distance indicators. The biggest galaxies in each cluster are of about the same size and luminosity. If this luminosity and size can be determined, they could be used to estimate the distances of clusters of galaxies. Figure 16-2 indicates how this method works. The clusters shown in each plate are at different distances from us; thus the galaxies in the Coma Berenices cluster, which is the closest of the three, appear larger than those in the other two clusters. The galaxies in the Hydra cluster appear the smallest, since it is the most distant. By measuring the apparent diameter of the largest galaxies in these clusters their distances have been estimated to be about 136 million light years for the Coma Berenices cluster, about 400 million light years for the Corona Borealis cluster, and about 1100 million light years for the Hydra cluster.

As astronomers seek to measure greater and greater distances they, of necessity, rely on those methods used to determine lesser distances. Although most distance measurements ultimately rely on the astronomical unit, some do not and consequently serve as a check. The

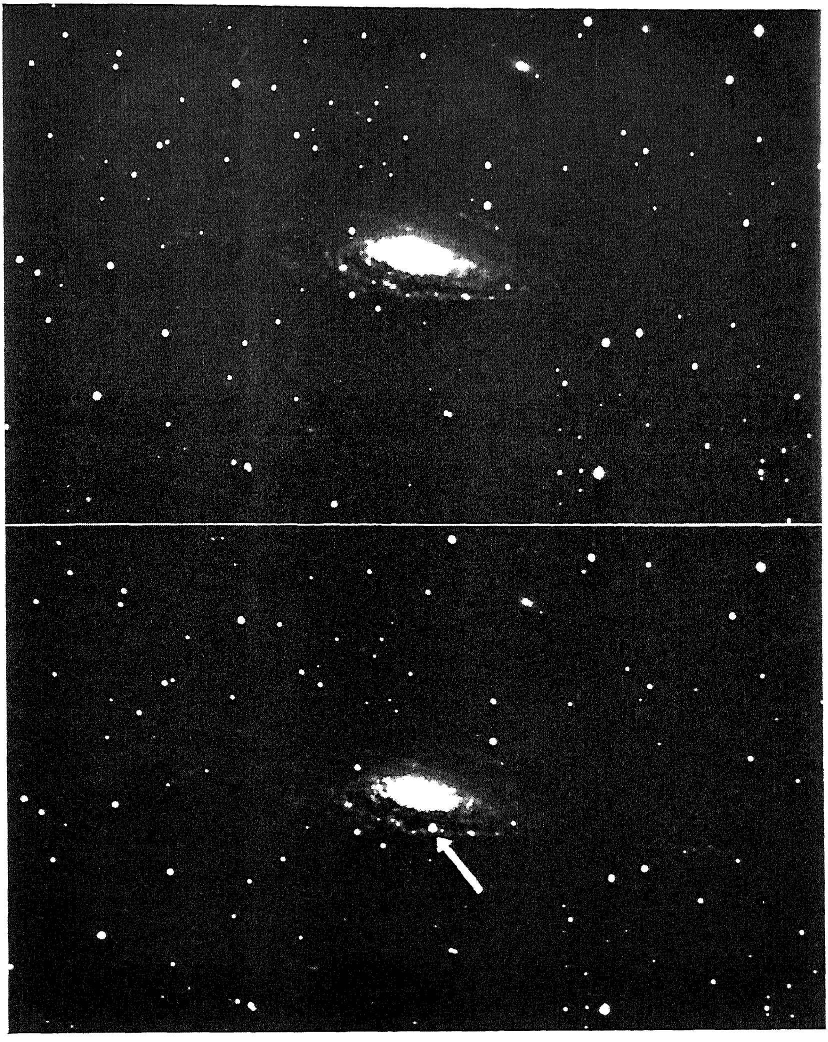


Figure 16-4 A supernova appeared in the galaxy NGC 7331 in 1959. The upper photograph was taken before the supernova erupted. (Lick Observatory)

distance of the Crab nebula (p. 293) is found by its rate of expansion, that of the Hyades cluster (p. 351f) by its rate of recession. But all distances outside our galaxy depend on the period-luminosity relation of the Cepheids. When it became clear that there are at least two types of Cepheids (see p. 278), the distances determined by means of Cepheids and consequently all intergalactic distances had to be increased by a factor of two, and the universe became twice as large.

In answer to the inevitable question, "Are the distances now accepted correct?" we can say only that these distances are the best values available and that they appear to be more nearly correct than the previously accepted ones. Before the distance scale was changed it was thought that the Andromeda galaxy was much closer and thus it (along with all the other galaxies) gave the impression of being smaller than it now appears to be. Astronomers felt uncomfortable about this, for it placed us in the unique position of belonging to the largest galaxy in the universe. But with the increase in the distance scale by a factor of 2, all the other galaxies have doubled their size. Andromeda is now larger than our own galaxy, and all the other galaxies are more nearly comparable in size to ours. Hence astronomers are more confident with the present distance scale than they were with the old one. As for our own galaxy its size was determined by RR Lyrae stars, not by Cepheids, and thus it was unaffected by the change in the distance scale. Now it is recognized that the E0 galaxies are the biggest of all.

The Red-Shift

There is yet another method of estimating distances of galaxies. In 1929 Hubble announced the results of studies which clearly indicated that with the exception of our Local Group *all* the galaxies have positive radial velocities, that is, a Doppler shift to the red. Furthermore, the farther the galaxy is from us the larger the red-shift. The problem of distance determination then became one of establishing the relationship between radial velocity and distance. This relationship can be estimated by plotting radial velocities against distances for galaxies whose distances have been determined by other methods. The resulting graph (Figure 16-5) is a straight line. Thus if the radial velocity of a galaxy is measured its distance can be estimated by consulting this graph. But as we shall see (p. 444), measurements of distances as large as a billion light years become questionable because of the very nature of their size.

We can see that the velocities on the vertical axis of the graph are considerably larger than any encountered in the discussions of other celestial objects. The largest red-shifts measured thus far amount to velocities of recession of about 151,000 miles per second, 81% *the velocity of light!*

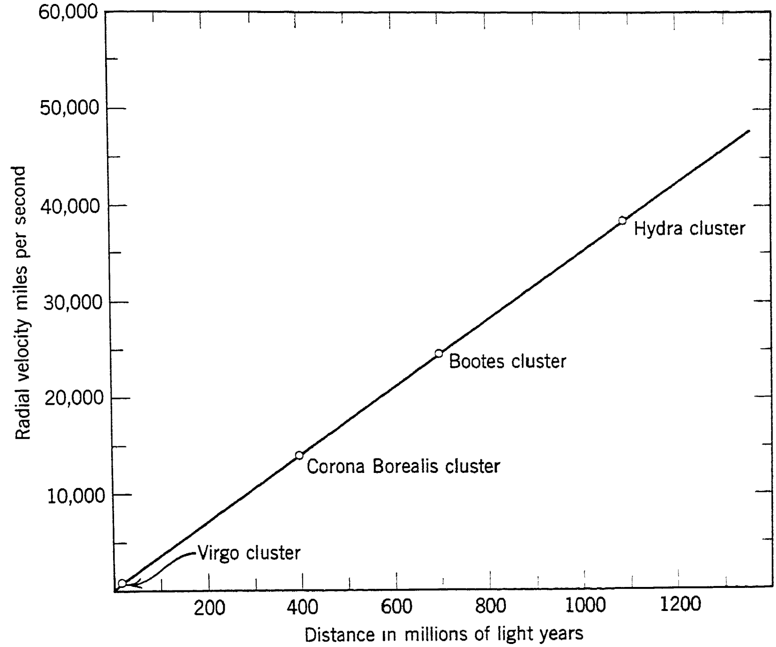


Figure 16-5 **The velocity of recession of galaxies increases with increasing distance from us. It must be remembered, however, that these distances represent the location of these clusters hundreds of millions of years ago.**

It appears that Hubble discovered a fundamental characteristic that may help man in his endeavor to understand the universe. Let us enumerate: (1) With the exception of our Local Group, *all* the galaxies have their spectra shifted to the red. If this is interpreted as a Doppler shift we are left with the conclusion that all these galaxies are receding from us! (2) The farther the galaxy is from us the larger the red-shift. Thus the velocity of recession increases with increasing distance. (3) These velocities of recession reach values interestingly close to the velocity of light.

Although it may be possible to interpret the red-shift as resulting from causes other than radial velocity, astronomers generally agree that the radial velocity interpretation is the most plausible. The natural consequence of this interpretation is that the universe is expanding. Thus the term red-shift has become synonymous with expansion of the universe.

The spectral lines normally used to measure the red-shift are the H and K lines of calcium, and the extent of their shift can be seen in Figure 16-6. The galaxy in the top spectrogram has a velocity of *only* 750 miles per second, and the lines have been shifted the distance shown by the arrow. In each succeeding spectrogram the H and K lines are shifted ever farther to the red until in the last plate they are shifted almost to the limit of sensitivity of the photographic emulsion used.

The fact that the red-shift increases with increasing distance from our galaxy places a maximum limit to the size of the observable universe. If galaxies exist at such distances that their radial velocities equal or exceed that of light, we will never be able to observe them because the light will never reach us. The largest radial velocity measured thus far is 81% the velocity of light.

Since all the galaxies are receding from us, we could infer that our galaxy is at the center of the universe, were it not that such an inference would place us only a step away from Plato, who put the Earth at the center of the universe. Astronomers are unwilling to accept any interpretation that places us in such a unique position. In general they prefer to base their thinking on the *cosmological principle: the appearance of the universe does not depend upon the observer's position.* That is, with the exception of local and small differences, the universe should appear the same to all observers no matter where they may be located—in our galaxy or in a very distant galaxy.

To explain the observed expansion on these grounds, the structure of the universe must be considered. Discussions of the structure of the universe, however, involve the theory of relativity expounded by Albert Einstein. Therefore it becomes necessary to consider certain aspects of that theory before proceeding with the structure of the universe.

The Principle of Special Relativity

At the turn of this century there were many observations conflicting with the physics derived since the time of Galileo. Yet that physics—now called *classical physics*—has proved extremely successful. For example, astronomers used Newtonian physics to predict the existence and location of an unseen planet, Neptune. Scientists still use it to place satellites in the orbit of their choosing, and to aim lunar probes at a particular location on the moon. It works. But it has its limitation.

In 1888 the famous Michelson-Morley experiment was performed which, when interpreted by Einstein, showed that the measured veloc-

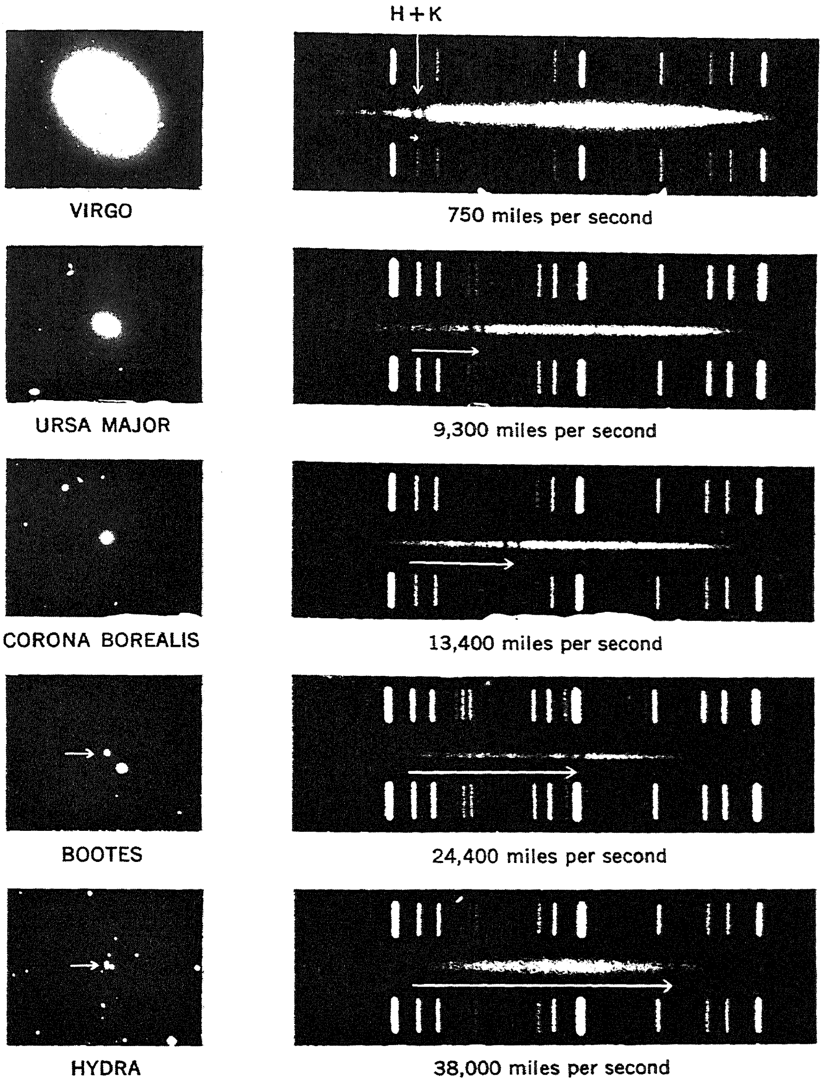


Figure 16-6 The actual spectrograms of five galaxies showing the extent of the shift of the H and K Lines of calcium to the longer wavelength. The apparent diameter of the galaxies decreases with increasing distance and they appear ever fainter. (Mount Wilson and Palomar Observatories)

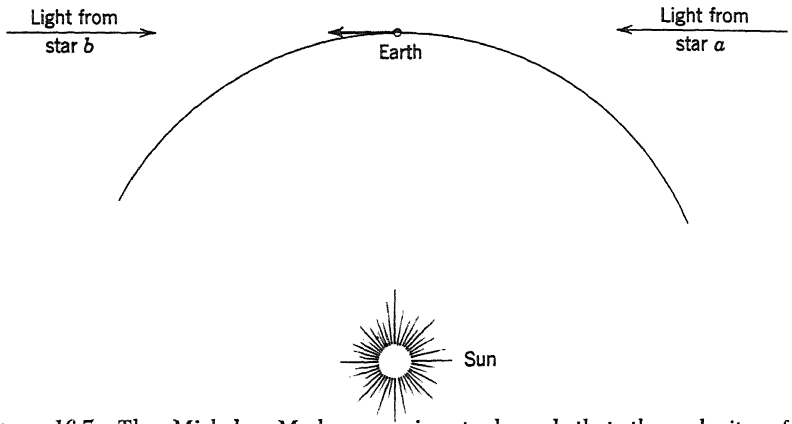


Figure 16-7 The Michelson-Morley experiment showed that the velocity of light as measured from the Earth of either star *a* or *b* is the same despite the motion of the Earth.

ity of light is independent of the motion of the observer. For example, if we were to measure the velocities of light emitted by star *a* (Figure 16-7) and by star *b*, we would find that those two velocities are equal! The measurement of each of those two velocities is independent of the motion of the Earth. This is not true of cars moving along a freeway; two cars moving at a velocity of 60 miles per hour but in opposite directions have a velocity with respect to each other of 120 miles per hour. But automobiles do not travel at speeds approaching that of light, and that is what makes the difference.

Stimulated to a large extent by the Michelson-Morley experiment, Einstein derived his *principle of special relativity* which describes the actions of systems moving relative to each other at *constant* velocities. The consequences of this principle are many, the concepts are profound and in conflict with our common sense, that is, the everyday garden variety of common sense.

The two basic postulates of special relativity are: (1) the velocity of light in free space is finite and constant for all observers, and (2) the laws of physics should be valid for all observers even if those observers move relative to one another. By adopting these two postulates Einstein showed that mass, length, and time change when transferred from one system to another if those two systems are moving at relativistic speeds (speeds close to that of light) with respect to each other.

If people outfit two spacecraft with essentially identical items and embark on a very rapid space flight such that the two spacecraft *A* and *B* travel at a relativistic speed with respect to each other, they will observe that objects on the *other* spacecraft appear both more massive and yet smaller—but smaller in only one direction, the direction of relative motion—than those once identical objects on their own spacecraft. They will find that the clocks on the *other* spacecraft will run slowly compared to the clocks on their craft. Passengers in both craft *A* and *B* will notice the same effects; the measurements of mass, length, and time depend upon the motion of the observer relative to that which he observes.

Ridiculous! says the skeptic. A kilogram is a kilogram, and a second of time is a second of time! But it has been proved beyond the shadow of a doubt, claims the atomic physicist. Electrons have been accelerated to speeds close to that of light and their mass does indeed increase; that is, a greater force is required to obtain the same acceleration to still greater speeds. Atomic particles called pions which serve as good clocks because they decay into other particles in a measurable duration of time have been accelerated to relativistic speeds and they decay more slowly!

In proposing these fundamental ideas Einstein did not give any reason for the apparent change in mass, length, and time when the observer moves relative to that which he observes. He simply showed conclusively that *these observations are the natural consequence of our living in a universe in which the velocity of light is finite and constant for all observers regardless of their motion.* The principles of special relativity have been proved valid and are the foundation for much of modern physics and astrophysics.

One of the many consequences of these fundamental concepts is the equivalence of energy and mass: $E = mc^2$. Recall that it is the conversion of mass to energy which permits astronomers to account successfully for the tremendous outpouring of energy from stars for billions of years. Another consequence of the postulates of special relativity is the fact that in free space no material object can travel with a velocity equal to or greater than that of light. This applies to electrons—this has been verified—and to galaxies which now have a measured velocity relative to the Earth of 81% the velocity of light! Are we approaching the end of the observable universe?

One of the nonrelativistic implications of the finite velocity of light is that astronomers really are cosmic historians, but today they *observe* history of the past.

Astronomers can only say what happened on the sun 8 minutes ago; they give the date of the supernova that resulted in the Crab nebula as A.D. 1054, even though the light from that explosion had already been traveling in space for about 4000 years before it reached the Earth. When they observe galaxies estimated to be 2 billion light years from the Earth, they see those galaxies as they were 2 billion years ago, not as they are "now." Therefore what use is it to say a certain galaxy is 2 billion light years away from us? None whatsoever, for all that an astronomer can say is that the galaxy's position 2 billion years ago is 2 billion light years away from our position at the present time. He cannot and never will be able to measure how far apart we are at any one given instant of time. There is no "now" in the universe because light has a finite velocity, and measurements of such vast distances become misleading.

This condition results from the fact that we cannot really separate time and distance as two separate and independent entities. If we wish to identify a certain event we must not only say where it occurred but when. If two people arrange to meet in the northwest corner of the 50th floor of the Empire State building, they have specified the three space coordinates: one along the north-south line, the other along the east-west line and the third along the up-down line. But they must also state at what time they will meet. Thus it takes four dimensions to identify a particular event, three of space and one of time, and these four dimensions are inextricably bound together.

Principle of General Relativity

Einstein also considered systems which move relative to each other with an accelerated motion. In doing this he evoked the *principle of equivalence: no instrument can determine any difference between the effects of gravity and the effects of acceleration.* As a thought experiment the reader should consider two spacecraft in "free space," that is, space free of any gravitational field. Let the first spacecraft *not accelerate;* its rocket motors are turned off. Without a gravitational field how will the occupants feel? Will they fall? No, of course not. The second spacecraft should *accelerate;* its rocket motors are turned on. How will the passengers of the accelerating spacecraft feel? Will they accelerate with the spacecraft? What force will act on them to accelerate them? After reasonable consideration the student should come to the conclusion that if the second spacecraft accelerates at

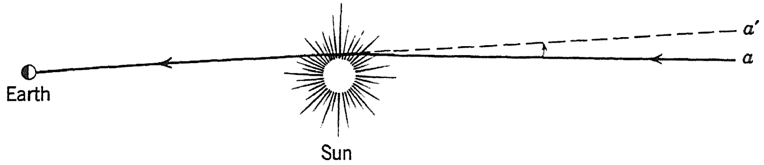


Figure 16-8 The light from a more distant star that passes close to the sun is affected by the sun's gravitational field in that the light bends toward the sun assuming a new direction of travel.

a rate of 32 ft/sec^2, the passengers would feel the same as if they were here on Earth. They would weigh the same; things would fall "down" the same as they do here on Earth. Down is toward that part of the spacecraft which exerts a force on and accelerates them. But, asked Einstein, what would happen to a beam of light shining at right angles to the direction of travel? He answered that question by showing that the beam of light would appear to bend! It would bend "downward" relative to the ship. So if a gravitational field is equivalent to an acceleration, a beam of light should bend downward in a gravitational field—a strong gravitational field.

Einstein predicted that light from a star which travels close to the surface of the sun will have its path deflected by the gravitational field of the sun (Figure 16-8). The apparent position of star *a′* would be radially displaced from its true position *a* in a direction away from the sun. The amount of this displacement increases with star light which passes ever closer to the edge of the sun. It has been observed in many total solar eclipses, for it is only at this time that stars so close to the sun's apparent disk can be seen. But this displacement is not a good quantitative verification of general relativity; the observed deflection is about 28% greater than the predictions of general relativity. No other theory has been proposed to replace the principles of general relativity, however, so we must rely on it as best we can.

The Structure of the Universe

In discussing the gravitational effect on light Einstein brought an entirely new concept to the attention of astronomers and cosmologists. If light travels unhampered by matter, it will travel in a straight line. Sometimes, however, a straight line is not a straight line as we commonly know it. If we define a straight line as the shortest distance between two points in space we have a definition that is completely general, but not so obvious.

To illustrate, let us consider the shortest distance *along the surface* of the Earth between San Francisco and London. This shortest distance is an arc of a *great circle*[*] on the Earth. An airplane pilot following this course must travel farther north than either San Francisco or London (see Figure 16-9).

But, you may argue, this course is not a straight line, it is an arc of a circle—which it is. Can you, however, distinguish between a straight line and the path of light through space? What constitutes a straight line for the beam of light as it passes close to the sun? That beam of light when viewed end-on appears as a point—the image of a star! Doesn't a straight line appear as a point when viewed end-on? Doesn't that beam of light continue to travel in a straight line even when near the sun?

Now we are confounded by geometry—but after all, geometry is a study of space, and space is a basic ingredient of the universe. If we define a straight line as the shortest distance between two points, we are no better off. It has been shown that light follows the shortest distance beween two points—when it does not pass from one medium to another (and the deflection of a star's light by the sun does *not* depend upon the solar corona). Einstein suggested that the universe may not be describable with Euclid's geometry. In fact, Einstein suggested that *a gravitational field curves space!* In curved space a straight line is not Euclidean, it becomes a *geodesic: the shortest distance between two points in space, curved or Euclidean.*

It is difficult enough to imagine "space" even without defining it, and an attempt to conceive of "curved space" seems at first quite hopeless. Space presumably is that which is between the atoms and dust particles in the interstellar regions. Space, then, being "emptiness," must be continuous; emptiness cannot be broken up into discrete particles with nothing between. So how can we possibly imagine space to be curved?

We can best imagine it if we reiterate what has been said—namely, that light travels along a geodesic in space, and if its path deviates from a straight line as defined by Euclidean geometry, the space must deviate in some way from Euclidean space.

Two intersecting straight lines in Euclidean geometry define a plane. A triangle, which in Euclidean geometry also defines a plane, has three angles the sum of which must add up to exactly 180°.

---

[*] A great circle is a circle whose center is coincident with the center of the sphere on which it is drawn.

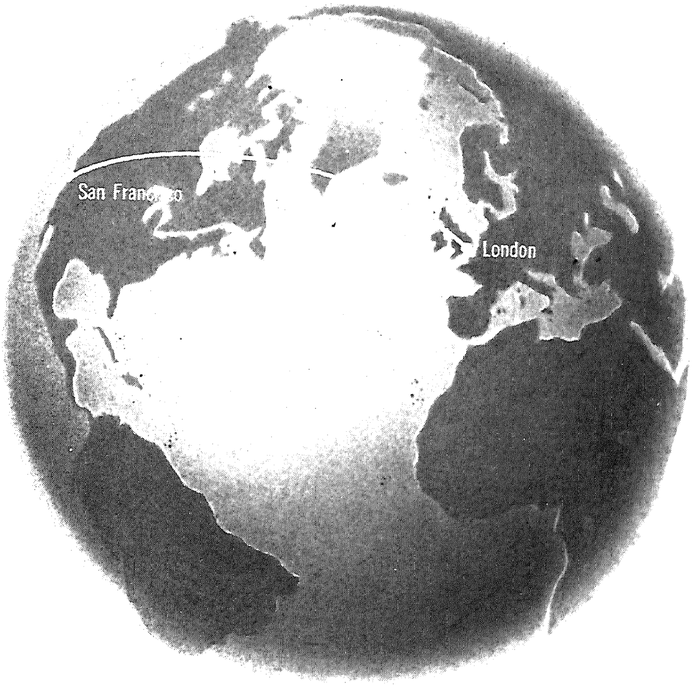


Figure 16-9 The great circle route from San Francisco to London.

But two geodesics in curved space do not define a Euclidean plane, and thus the three angles of a triangle in curved space do not add up to 180°. The angles of a *spherical triangle* must exceed 180° but be *less than* 540° (Figure 16-10).

If the mass of the sun is enough to curve space in its immediate vicinity, perhaps all the mass in the universe is enough to curve all the space in the universe. Maybe a geodesic in the universe is not a Euclidean straight line. Perhaps Euclidean geometry is so fundamental to us only because our everyday straight lines and planes are so extremely small when compared to the universe. A ship crossing the Pacific Ocean seems to be continually traveling in a straight line, yet we know that it is traveling on a spherical Earth and thus it is actually traveling along a Euclidean circle. The world was once thought to be flat; it has been proved to be nearly spherical. Our universe can be thought of as being a Euclidean universe, but it is possible that it is not.

To illustrate a non-Euclidean universe we must use our imagination.

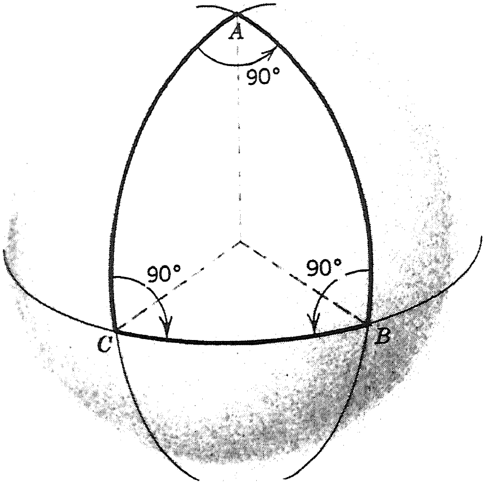
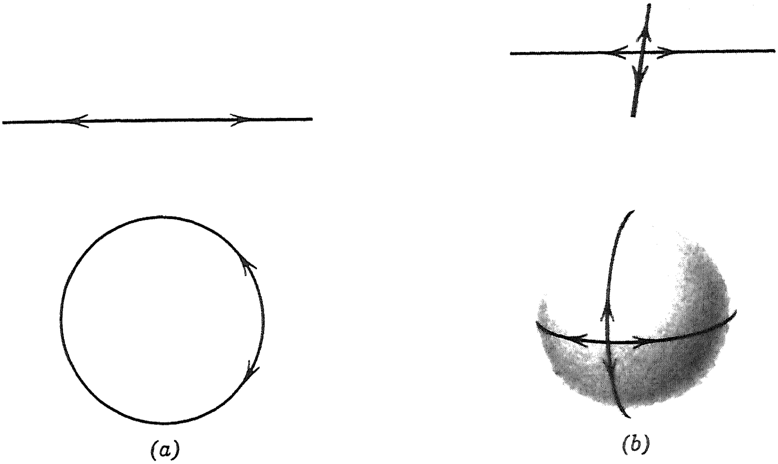


Figure 16-10 The three angles of this spherical triangle add up to 270°.

Let us imagine a bug that has only one dimension—a creature that can move and conceive of motion only along a straight line, or only along the circumference of a circle (Figure 16-11*a*). If this bug resides in a straight line, he can travel forward or backward for an infinite distance and not retrace his steps, but he cannot move to the right or left. If he lives on the circumference of a circle he can still travel forward or backward an infinite distance. In order to travel an infinite distance he must first complete one trip around the circle and then retrace his steps for an infinite number of revolutions. He cannot travel along a radius of the circle; in fact he cannot even imagine a radius for he can conceive of only one dimension—the forward and backward directions.

Now let us imagine a two-dimensional creature—a beetle (Figure 16-11*b*). He can reside in a Euclidean plane and move either forward or backward an infinite distance, or to the right or left an infinite distance without retracing his steps. On the other hand, if he lives on the surface of a sphere, he can travel either forward or backward, or to the right or left an infinite distance, but he will eventually return to his starting place and retrace his steps. He cannot travel or even conceive of motion along the radius of the sphere.

A creature living in a straight line or in a plane would live in a Euclidean universe of one and two dimensions respectively. A crea-



(a) (b)

Figure 16-11 (*a*) A one-dimensional bug can travel along a line or along a circumference of a circle. (*b*) A two-dimensional beetle can travel along a plane or along the surface of a sphere.

ture living on the circumference of a circle or on the surface of a sphere would live in a non-Euclidean universe of one and two dimensions respectively. Geodesics in these non-Euclidean universes are not Euclidean straight lines.

Do we, as three-dimensional creatures, live in a universe that is Euclidean and thus one step up from the line-plane sequence? Or do we live in a non-Euclidean universe that is one step up from the circle-sphere sequence? If our universe is Euclidean it could be infinite in extent; we could imagine ourselves traveling in space along a straight line for an infinite distance and never retracing our steps. If we live in a non-Euclidean universe which is the third in the sequence circle-sphere-universe, we could travel along a geodesic for an infinite distance but would eventually return to our starting place and begin to retrace our steps.

A Euclidean universe is easier to imagine, for it would be three-dimensional and infinite in extent. A straight line would be a Euclidean straight line. A non-Euclidean universe of three dimensions is *impossible* to imagine. The two-dimensional beetle in the surface of his spherical universe cannot possibly conceive of a third dimension, yet the radius of his sphere is in that third dimension. It is a line that is perpendicular to the only two directions the beetle can imagine. If our universe is of this type, it has a radius that is in the fourth

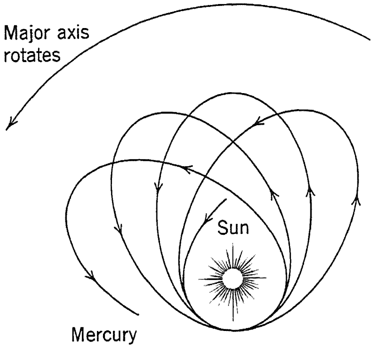


Figure 16-12 The major axis of Mercury's orbit rotates according to Einstein's theory of relativity and not according to Newton's law of gravitation.

dimension and is mutually perpendicular to the three directions we can imagine. Obviously we cannot conceive of this fourth dimension, but we can treat it mathematically; we can study its properties by extending our observations from a circle to a sphere and from there to this type of non-Euclidean universe.

Einstein treated the universe as a four-dimensional space-time continuum. The Earth travels in a region of the space-time continuum that is affected by the gravitational field of the sun. It therefore travels along a geodesic through this space-time continuum. Such concepts, although difficult to imagine, have been verified by the orbit of Mercury. Newton's law of gravitation is not able to account for the orbit of Mercury in that the major axis of its orbit rotates more rapidly than gravitational theory predicts. Mercury moves in a manner shown in the exaggerated drawing in Figure 16-12. Einstein's concept of the universe is mathematically able to account for this peculiar motion, which Mercury displays more than any other planet because it is so close to the sun.

How can we determine the shape of our universe? To answer this question let us again resort to the two-dimensional analogy. We can place dots on a plane in a random fashion but with an overall uniform density and let these dots represent clusters of galaxies in a Euclidean two-dimensional universe (Figure 16-13). Assume that dot *A* represents the cluster of galaxies in which our two-dimensional creature exists. If we count the clusters of galaxies in ever-increasing circles about him we find that their number increases as the square of the distance away from him. There are 3 dots in the first circle whose

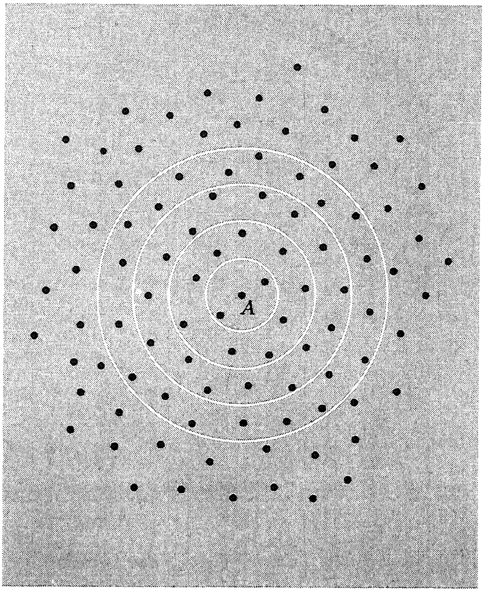
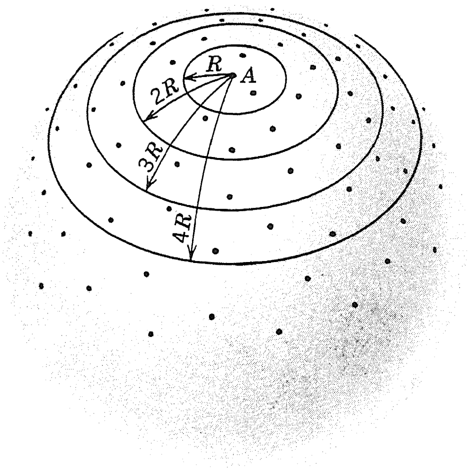


Figure 16-13 In Euclidean space, the number of randomly spread dots included within ever larger circles increases as the square of the radii for two-dimensional space and as the cube of the radii of spheres for three-dimensional space.

radius is 1 unit, 12 dots (3×2^2) in the circle of radius 2 units, 27 (3×3^2) in the circle of radius 3 units, etc.

Now let us represent a non-Euclidean two-dimensional universe by placing dots on a sphere (such as ink dots on a balloon) and label one of them *A* (Figure 16-14*a*). If our beetle now counts these dots about him he will find that their number will increase less rapidly than the square of the distance from him. There may be 3 dots in the first circle but only 11 in the second and 25 in the third. The reason for this is that there is less area on the surface of a sphere of a given radius than there is in a plane within a circle of the same radius, as may be seen by taking one half of a grapefruit peel and flattening it out on a table. The grapefruit peel will split in several places. This shows that there is less area on the surface of the half grapefruit peel than in a circle on a plane with a radius equal to one-fourth the circumference of the grapefruit (Figure 16-14*b*).

It is also conceivable that non-Euclidean two-dimensional space could exist in which the number of dots in ever-increasing circles



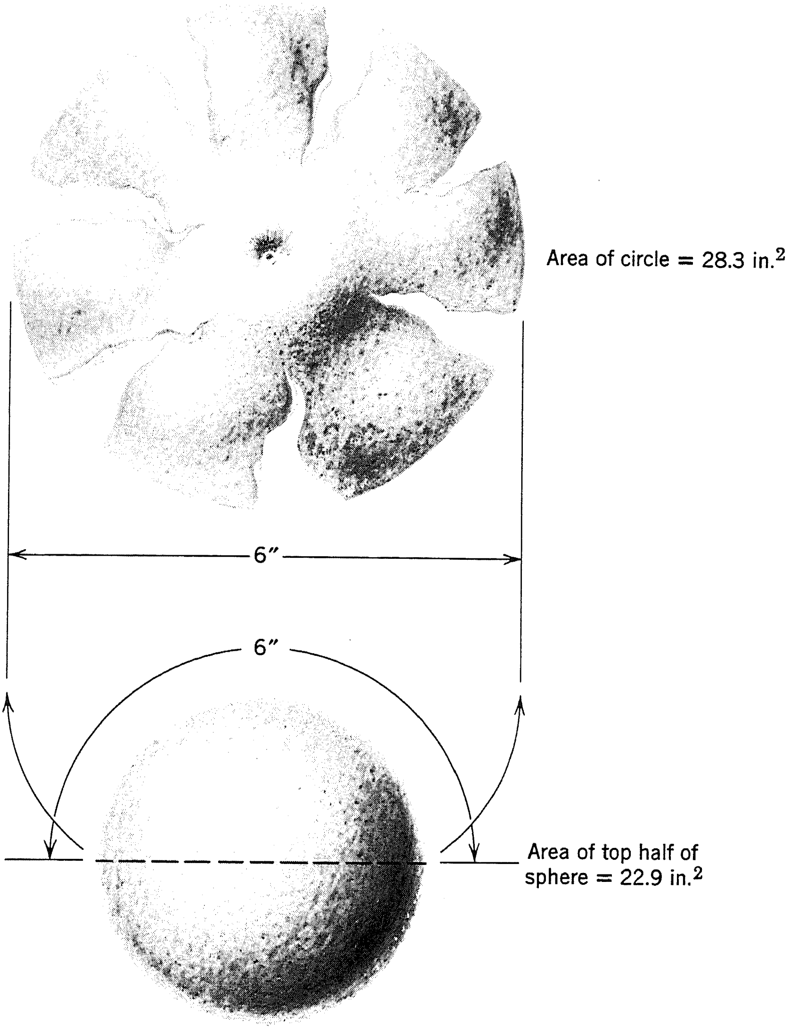
(a)

Figure 16-14 (*a*) In positively curved space, the number of randomly spread dots included within ever larger circles increases less rapidly than the square of the radii for two-dimensional space and less rapidly than the cubes of the radii of spheres for three-dimensional space. (*b*) There is more area included within a circle of given radius if that circle is on a plane than if the circle is on a sphere with the radius along the circumference of that sphere.

would increase more rapidly than the square of the distance out. Such space would take the shape of a western saddle (Figure 16-15*a*). It is not difficult to imagine a flimsy saddle that could be flattened out on the floor of the saddle room and in so doing the saddle would have to fold in several places in order to lie flat on the floor (Figure 16-15*b*). Thus there is more area on a saddle within a circle of given radius than on a plane within a circle of the same radius.

Space that follows a form similar to the Euclidean plane is called *zero-curvature* space; that which appears on the surface of a sphere is called *positively curved* space; and the space on a saddle is *negatively curved* space.

To determine the shape of our universe we must count clusters of galaxies in ever-increasing *volumes* of space, since our space is three-dimensional. Therefore if the number of clusters increases as the *cube* of the distance from us, our universe is Euclidean. If the number of clusters increases less rapidly than the cube of the distance from us, our universe has a positive curvature (circle-sphere). If the number of clusters of galaxies increases more rapidly than the


Area of circle = 28.3 in.²
6″
6″
Area of top half of sphere = 22.9 in.²

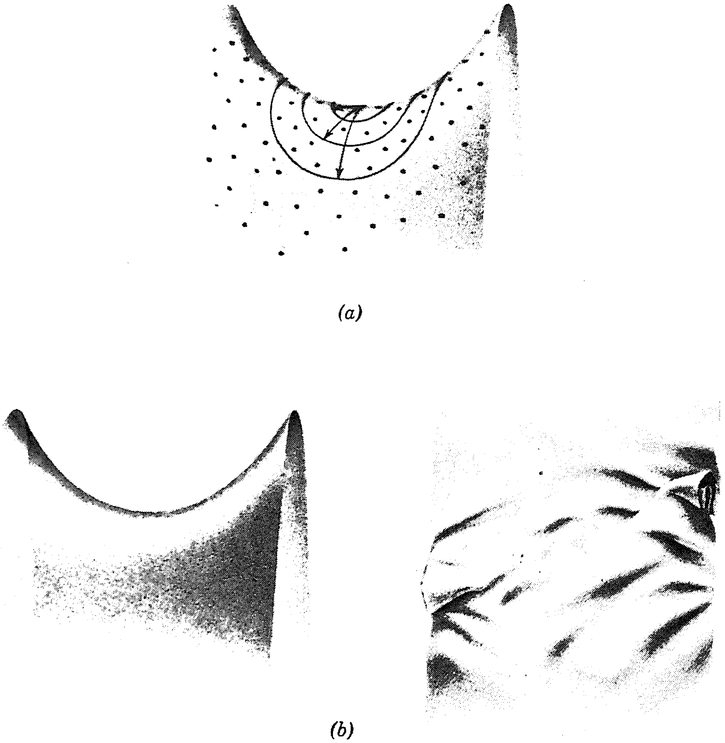
(b)

Figure 16-14 (Continued)

cube of the distance from us, our universe has a negative curvature (saddle). Such counts have been made, but uncertainties in distance and thus in time have so far prevented them from becoming valid tests.

Expansion of the Universe

Expansion can be explained in both a Euclidean and a non-Euclidean universe. If the universe is Euclidean, we must imagine that



(a)

(b)

Figure 16-15*a* In negatively curved space the number of randomly spread dots included within ever larger circles increases more rapidly than the square of the radii for two-dimensional space and more rapidly than the cubes of the radii of spheres for three-dimensional space. (*b*) There is less area included within a circle of given radius if that circle is on a plane than if the circle is on a saddle with the radius following the curvature of that saddle.

the space between the clusters of galaxies becomes larger as the universe expands, just as the space between the raisins in bread becomes larger as the bread is baked. The clusters themselves do not become larger but the space between them increases. The larger the distance between any two clusters the more rapidly the clusters recede from one another. Thus the rate of expansion increases with increasing distance between clusters of galaxies. The effect would appear the same from anywhere in the universe.

If the universe is non-Euclidean, we can still account for the observed effects of expansion. Let us recall the two-dimensional beetles on the surface of a balloon on which dots represent the clusters of galaxies. If the balloon is further inflated the distance between any two dots will increase. Dots farther apart recede from each other more rapidly than those close together, for there is more rubber between these dots to stretch. Again the rate of expansion increases with increasing distance and the effect is the same no matter from which dot the observations are made. This analogy fails in one respect—the dots on an inflating balloon will become larger whereas the clusters of galaxies remain the same size.

The Origin of the Universe

The shape of our universe, whatever it may be, has not yet been determined, but it must be related in some way to the origin of the universe. Did our universe have a beginning? If so, how did it begin? And when?

There are currently two main theories concerning the evolution of the universe. One is called the *evolutionary theory,* for it assumes a definite beginning at a definite time. The other, the *steady-state theory,* maintains that the universe has an infinite age and thus no beginning or ending.

According to the evolutionary theory all the matter in our universe was once contained in a *primeval nucleus,* the density of which was perhaps equal to the density of the nucleus of an atom. This primeval nucleus exploded. One hour after it exploded, it had a temperature of 250 million degrees Kelvin. As the universe has continued to expand it has continued to cool to its present over-all temperature of about 50°K (−370°F).

By determining the rate of expansion and the size of the present universe it is possible to retrace the process to find the date of the explosion. The date considered at present is roughly 13 billion years ago. Presumably the rate of expansion decreased as the universe expanded just as the velocity of a rock thrown in the air decreases. The rate of decrease depends on the size of the universe and its mass, for as the universe expands the kinetic energy of the galaxies is converted into potential energy.

If Newton's law of gravity holds in such a universe, there are basically two possible modes of expansion. If the universe is expanding fast enough, the clusters of galaxies exceed the escape velocity of the universe and will continue to recede from each other forever. The space density of clusters of galaxies then continues to decrease until in some distant epoch the only galaxies visible from the Earth (if it should still exist then) would be those in our own cluster. If, however, the universe is not expanding at such a "breakneck" speed, eventually the mutual gravitational forces of all the galaxies will stop the expansion and reverse the process. The universe will start to contract and eventually return to its primeval nucleus. Presumably it could expand again, repeating the process over and over, becoming a *pulsating universe.*

The theory of a pulsating universe implies that the universe has many beginnings, for it would start anew every time it expanded from the primeval nucleus. But the pulsations could have been going on for an infinite time, and we could never determine this, for each time it contracts into the primeval nucleus all traces of its past life must be completely erased forever. The best we can do would be to determine how long it has been since the most recent primeval nucleus.

There is also the possibility that the universe started as an infinitely thin gas from which it contracted into the primeval nucleus. It then expanded, but on this theory the rate of expansion would be too great to allow it to contract again and it would continue to expand for an infinite length of time.

The idea of the steady-state universe results from an extension of the cosmological principle. Proponents of this theory make the point that Einstein showed that the universe is a four-dimensional space-time continuum. Therefore, the cosmological principle should be restated: *the appearance of the universe does not depend upon the observer's position in either space or time.* That is, with the exception of local and small differences, the universe should appear the same to all observers no matter where they may be located or when—at what astronomical epoch—they should observe it. The universe should appear the same, in its over-all aspects, if viewed 5 billion years ago or 5 billion years from now. Surely some stars will have been born and died out; individual galaxies will have changed, but the over-all aspects of the universe should remain constant throughout all space and time.

The steady-state theory assumes that the universe never began, or to be more precise, that it is in a process of continual creation. As the universe expands, its over-all density decreases, but there is a universal lower limit to this density, below which it cannot exist. As the density approaches this lower limit, more matter is created to increase the density once more. Thus as the universe continues to expand, new matter is created to fill the void. The newly formed matter is hydrogen that eventually forms clusters of galaxies. Each new cluster of galaxies lives out its life as the universe continues to expand, while new clusters of galaxies form. New galaxies form, and old ones die, but the universe is always at the same density and there are always galaxies of different ages. Thus the universe will be the same if examined at any epoch. Although individual galaxies and clusters will have changed, the over-all picture will not have changed. This is what is meant by a "steady-state" universe that is not only infinite in age but infinite in extent.

Which theory, if either, is correct, is difficult to tell. Future observations should reveal the answer. If we can find very old galaxies and very young galaxies, the steady-state theory will gain favor. However, the number of old galaxies in any one part of the universe should be rather small, since they are diluted by newer galaxies as the universe continues to expand.

Another observation that might help to reveal the answer lies in the rate of expansion of the galaxies as their distance from us increases. If the steady-state theory is correct the rate of expansion should increase steadily with increasing distance, for the universe, by and large, is the same throughout all space and time. If, however, the rate of expansion should increase more rapidly, it would support the evolutionary theory, for as we see ever more distant galaxies we also see back in time and thus observe the rate of expansion of the universe as it existed some one billion years ago or more. Let us recall that according to the evolutionary theory the rate of expansion should decrease with increasing time; as we look *back* in time, then, this rate should increase.

Still another critical test is the observation of the density of galaxies in the universe which, according to the steady-state theory, is constant over all time and space. The evolutionist, on the other hand, says that the density of galaxies in space is continually decreasing with time; thus as we look *back* in time by observing ever more distant galaxies, the density should increase. The most distant observable galaxies are perhaps 5 to 10 billion light years away; according to the evolutionary theory these should be about 5 to 10 billion years younger than those closest to us and consequently in an earlier stage of development. The steady-state theory, on the other hand, says that since the universe is the same over all time and space, the very distant galaxies should be similar to those closest to us.

Quasars

The correlation of discrete radio sources with images on photographic plates has led astronomers down many new paths. Exploding galaxies was one such path. Another path is the very puzzling objects which have been dubbed *quasars,* short for quasi-stellar sources. (When an object is discovered it must be named for purposes of classification and conversation, but its name may have nothing to do with its actual structure—since its structure may not be known when the name is applied. For example, quasi-stellar source means essentially the same thing as asteroid! But there the similarity ends.)

Quasars were observed first as discrete radio sources and so individuals are named according to their number in the *Third Cambridge* catalog: 3C 48, 3C 273, etc. They were later identified as bluish "stars" (Figure 16-16), with unrecognizable spectral lines. Martin Schmidt of the Mount Wilson and Palomar Observatories, however, showed that these strange spectral lines are really lines of the far ultraviolet displaced into the near ultraviolet and visible part of the spectrum. Quasars suffer a severe case of red-shift.

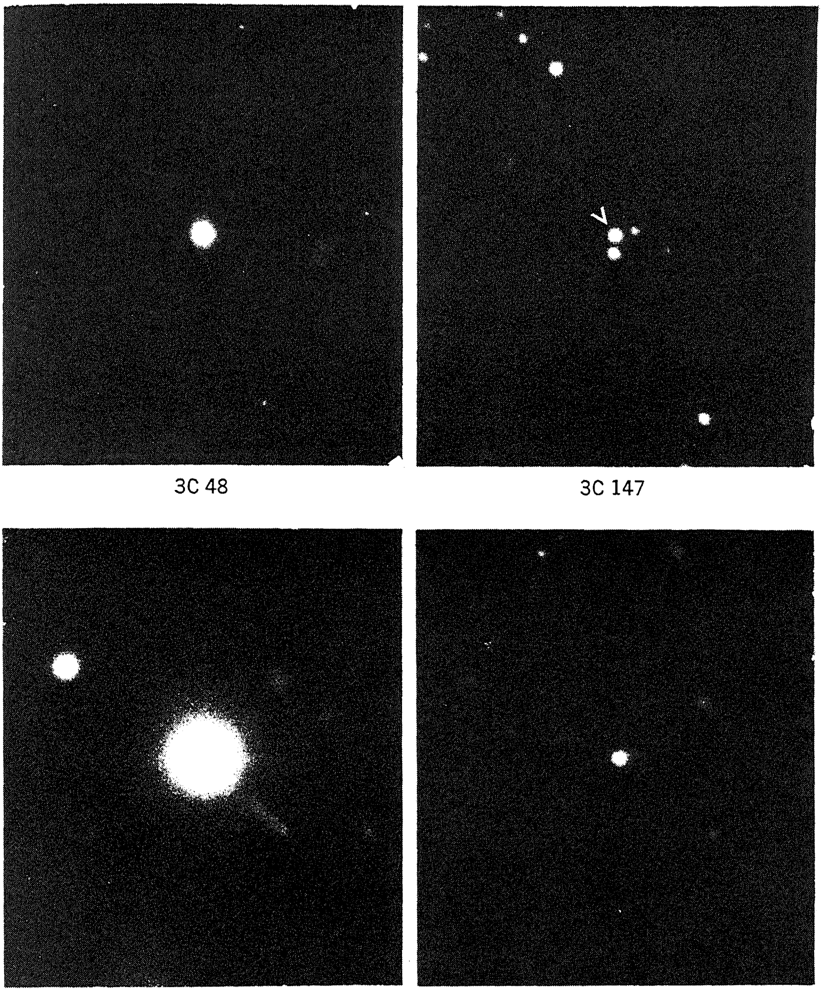
The red-shift of 3C 9 is so large that the change in wavelength exceeds the unshifted wavelength by a factor of 2! This means that the Lyman-α line of hydrogen with an unshifted wavelength of 1216 Å has an apparent wavelength of 3638 Å ($1216 + 2 \times 1216$). The Balmer $H\alpha$ line with an unshifted wavelength of 6563 Å appears in the infrared region at 19,689 Å. To translate this into a radial velocity using the classical relationship

$$\frac{\Delta\lambda}{\lambda} = \frac{v}{c}$$

would yield a velocity of recession twice that of light! How can we see it? We must, therefore, use the relativistic Doppler shift which gives the ratio of the change in wavelength $\Delta\lambda$ to the unshifted wavelength λ_0 as

$$\frac{\Delta\lambda}{\lambda_0} = \sqrt{\frac{c + v}{c - v}} - 1$$

where v is the relative velocity of the source and c the velocity of light. But one wonders at this point what it means to speak of astronomical objects with velocities of recession 80% the speed of light. Can these objects be completely understood in the framework of our present ideas of the universe?



3C 48

3C 147

3C 273

3C 196

Figure 16-16 Four different quasars. The jet can be seen extending from 3C 273. (Mount Wilson and Palomar Observatories)

As these fantastic red-shifts were observed astronomers sought other observations. Some of the quasars are fairly easy to observe optically; 3C 273 has an absolute magnitude of 12. A good 8-in. amateur telescope will enable one to see it. Consequently, it has been recorded on thousands of patrol plates taken since the late 1880's at a number of observatories. By examining these plates it has been determined that the brightness of 3C 273 varies with a slow oscillation of about

½ magnitude over a period of roughly 12 to 13 years. But there are more rapid irregular variations in brightness, too. Observations show that the radio brightness of 3C 273 has *increased* more or less steadily by about 40% in the first three years of observations. Other quasars have *decreased* in radio brightness over the same interval of time, and still others have remained fairly constant.

The occultation of 3C 273 by the moon enabled astronomers to identify two radio sources. The photograph (Figure 16-16) reveals a jet-like appendage which supplies some of the energy, the remainder comes from the main part of the object itself. Other quasars appear as single sources of radio energy. Still other objects with quasar-like spectra do not emit enough radio energy to be detected with present equipment.

If the tremendous red-shift of the quasars is interpreted to be part of the expansion of the universe as described by Hubble it means that these are the most distant of all objects observed. According to the revised Hubble relationship, 3C 9 would have to be 10 billion light years away! But what does that mean? Is Hubble's relationship reliable at such great distances? Quasar 3C 273 would be 2 billion light years away—and yet it is a twelfth-magnitude object visible in a small telescope! If we assume it to be at this great distance, then in order to be seen so brightly, it must emit energy at the rate of 10^{46} ergs per second! The sun emits 10^{33} ergs per second, so 3C 273 must emit energy at the rate of 10,000 billion suns! The numbers stagger the imagination.

If quasars are not so lavish in their outpouring of energy, they must be closer to us. Perhaps they are stars in our own galaxy moving with prodigious speeds. But they all, without exception, exhibit red-shifts; there is not a blue-shift in the lot. Not one of them exhibits any proper motion, which they should were they so close to us and moving so fast.

Perhaps general relativity can help us. Light emitted by a source with a very strong gravitational field will lose energy in escaping from that source. By losing energy, the wavelength increases, so the entire spectrum shifts to the red. This is called the *gravitational red-shift*, and has been observed in white dwarf stars which are members of binary star systems. Its existence, therefore, is established, even if precise measurements of the extent of gravitational red-shift have yet to be made. Perhaps quasars are extremely dense stars, so dense that the gravitational fields are strong enough to cause the observed red-shift. But the spectra of quasars have forbidden lines of ionized neon and oxygen, and in objects so dense as to produce a gravitational red-shift the gases must be too dense to permit elements to emit the same forbidden lines seen in rarified gaseous nebulae.

One piece of evidence, the radio spectral line at 21 cm, established an extragalactic distance of 3C 273. That this quasar lies *beyond* the great Virgo cluster of galaxies is demonstrated by the observation that some of the intergalactic neutral hydrogen of that cluster has caused some absorption of the 21-cm line of 3C 273. The Virgo cluster is about 36 million light years away, so astronomers are left with the problem of accounting for tremendous outpourings of energy.

The true nature of quasars remains a puzzle. It is possible that they are elliptical galaxies in the process of formation. They may have evolved to the extent that the first blue super-giants to form have evolved into supernovae. If 50 billion stars had formed and if one in 500 is a blue super-giant, then 100 million supernovae would erupt in a relatively short interval of time, releasing a lot of energy.

It is possible that quasars may result from a gravitational collapse. General relativity predicts that a very massive object with a high density could collapse in on itself. In such a condition the gravitational potential energy is converted into electromagnetic energy. The quasar would half-destroy itself before collapsing out of sight. It would collapse out of sight when the gravitational field became so strong that no light was able to leave—when the escape velocity reached the velocity of light! It could then be observed only by its gravitational field.

But having learned something of the changes that observations have forced man to make in his thinking of the universe we should be ready to expect an explanation of quasars which stems not from the present ideas of physics and astronomy, but perhaps from some fundamental mechanism not yet imagined by man. Startling observations demand bold explanations; new ideas stimulate new observations. This is the way science progresses.

Basic Vocabulary for Subsequent Reading

Cosmogony
Cosmology
Geodesic
Negatively curved space
Positively curved space
Quasars
The red-shift
Zero curvature space

For Further Reading

Barnett, Lincoln, *The Universe and Dr. Einstein,* William Morrow and Co., New York, 1957.

Bonner, William, *The Mystery of the Expanding Universe,* Crowell Collier and Macmillan, New York, 1964.

Hoyle, Fred, *Galaxies, Nuclei, and Quasars,* Harper and Row, New York, 1965.

Messell, H. and S. T. Butler, *The Universe and its Origin,* St. Martin's Press, New York, 1964.

Munitz, Milton K., *Theories of the Universe,* The Free Press, New York, 1957.

Chiu, Hong-Yee, "Gravitational Collapse," *Physics Today,* p. 21, May 1964.

Gamow, George, "Gravity," *Scientific American,* p. 94, March 1961.

Green, Louis C., "Dallas Conference on Super Radio Sources," *Sky and Telescope,* p. 80, February 1964.

Green, Louis C., "Observational Aspects of Cosmology," *Sky and Telescope,* p. 199, April 1966.

Green, Louis C., "Relativistic Astrophysics," *Sky and Telescope,* part I, p. 145, March 1965; part II, p. 226, April 1965.

Greenstein, Jesse L., "Quasi-Stellar Radio Sources," *Scientific American,* p. 54, December 1963.

Robinson, I., et al., "Relativistic Astrophysics—A Report on the Second Texas Symposium," *Physics Today,* p. 17, July 1965.

Schmeidler, F., "The Einstein Shift—An Unsettled Problem," *Sky and Telescope,* p. 217, April 1964.

"The Local Group of Galaxies," *Sky and Telescope,* p. 21, January 1964.

"Quasar Hypotheses," *Scientific American,* part I, p. 38, August 1964; part II, p. 50, February 1966.

"Quasi-Stellar Galaxies," *Sky and Telescope,* p. 67, August 1965.

"Variable Radio Emission from Quasars," *Sky and Telescope,* p. 277, November 1965.

"Where Did Cosmic Helium Originate?" *Sky and Telescope,* p. 88, February 1965.

Questions and Problems

Chapter 1

1. England was late in adopting the Gregorian calendar. What trouble did this cause the British government?
2. Locate Polaris, and by watching the stars during the night verify the statements that each travels in a circle concentric with the North Celestial Pole.
3. With a little effort, a great deal of enjoyment can be gained by learning to recognize the more important constellations.

Chapter 2

1. Write the following numbers in the exponential system:
   (*a*) 93,000,000
   (*b*) 24,000,000,000,000
   (*c*) 0.034
   (*d*) 0.000 000 45
2. Write the following wavelengths in Ångstroms:
   (*a*) 4.861×10^{-5} cm
   (*b*) 6.0×10^{-3} cm
   (*c*) 3.4×10^{-7} cm
   (*d*) 21 cm
3. If the wavelength of a spectral line is 4,500 Å, what is its frequency (the velocity of light $c = 3 \times 10^{10}$ cm/sec)?
4. Determine the Doppler shift for the hydrogen spectral line (λ = 4,860 Å) for stars whose radial velocity is:
   (*a*) 10 miles per second
   (*b*) 100 miles per second
   (*c*) 1,000 miles per second
5. Determine the radial velocity of an object if the K-line of ionized calcium (λ_0 = 3,930 Å) is shifted so that it has a measured wavelength of 4,930 Å. Is the object receding from or approaching the Earth?
6. In question 5, the K-line was shifted by 1,000 Å. How far would the 21-cm line be shifted if radio observations were made of the same object?

7. What is the resolving power of:
    (*a*) the 200-in. telescope at a wavelength of 5,000 Å?
    (*b*) the 1,000-ft radio telescope in Puerto Rico at a wavelength of 21 cm?

Chapter 3

1. A brick is first lifted from the floor to the table top, and then released to fall back to the floor again. Describe its changes in energy.
2. What is the kinetic energy of the Earth if its mass is 6.0×10^{27} grams and its velocity about the sun is 3.0×10^{6} cm/second? (The unit of energy is the *erg*.)
3. (*a*) Explain three ways in which energy can be given an orbital electron.
    (*b*) What is the natural consequence of the orbital electron's absorbing this energy?
    (*c*) By what single way can the electron lose this added energy?
4. Explain the process of photoionization.
5. Describe the cause of spectral-line pressure broadening.
6. What are the limitations of electron transitions within a given atom?
7. What are the limitations of energy emitted during the capture of a free electron?

Chapter 4

1. Draw, to scale, the sun and its nine planets.
2. Explain how the period of rotation of a planet can be determined by the use of radar.
3. Compare the surface of Mars with that of the Earth and that of the Moon.
4. Describe the radio signals received from Jupiter.
5. What are the significant differences between the terrestrial planets and the Jovian planets?
6. Discuss the possibility of life on each of the other planets of our solar system.
7. Explain how the period of rotation of Pluto was determined.

Chapter 5

1. Explain why Mercury and Venus are always seen relatively close to the sun.
2. If the sun moves east relative to the stars, why does it set in the west?

3. Draw the planetary distances from the sun on a scale such that the radius of Mercury's orbit is ¼ in. Can you draw the planets to the same scale?
4. Describe two proofs of rotation and two of revolution of the Earth.
5. Explain why the Earth has seasons.
6. If a planet revolved about the sun in an orbit with a radius vector of 2.8 A.U. what would be its period of revolution?
7. Describe the appearance of the sky both day and night through each of the four seasons as seen by (*a*) explorers on the North Pole; (*b*) people on the equator.
8. Calculate the gravitational force of attraction (in dynes, 1 dyne = 2.2×10^{-6} pound) between the Earth and the moon from the following information:

   mass of the Earth = 6.0×10^{27} grams
   mass of the moon = 7.3×10^{25} grams
   radius of moon's orbit = 3.8×10^{10} cm
   G, gravitational constant = 6.7×10^{-8} dyne-cm^2/gram^2

9. Trace the ecliptic in the sky by following the moon and as many planets as you can see.
10. Plot the location (right ascension and declination) of Mars throughout the year from its positions as listed in the current *American Ephemeris* and *Nautical Almanac*. Its motion can readily be seen by plotting positions for the 10th, 20th, and the last day of every month.
11. An artificial satellite was launched with a period of 100 minutes. By using the moon as a comparison find the radius of the artificial satellite's orbit. (Radius of moon's orbit is 2.4×10^5 miles; its period is 2.4×10^6 seconds.) Now find the velocity of this artificial satellite in its orbit.
12. Explain why there is both a sidereal and a solar day. Which is longer and why?

Chapter 6

1. Draw the Earth-moon system to scale and indicate the center of mass of the system.
2. After the next new moon follow the moon from night to night through at least one complete cycle of phases (weather permitting).
3. Describe the appearance and possible origin of three features of the lunar landscape.
4. Explain the relationship between solar eclipses, lunar eclipses and the lunar nodes.

Chapter 7

1. Describe how the height of a meteor can be determined.
2. Explain one observation which indicates that meteors are not much larger than a grain of sand.
3. What evidence indicates that meteorites may be small minor planets swept up by the Earth?
4. What evidence indicates that many meteorites may have belonged to only two bodies?
5. Describe the appearance of a comet as it approaches the sun from beyond the orbit of Neptune, passes close to the sun, and returns to aphelion again.
6. What information leads astronomers to suspect that meteor showers are related to comets?

Chapter 8

1. As an example of the methods of age determination, 3 cubic cm of water are poured into a large graduate cylinder every 15 seconds (analogous to the accumulation of lead from the radioactive decay of uranium). After some time it is determined that the graduate cylinder contains 57 cubic cm of water.
   (*a*) Assuming that the graduate cylinder was empty before the first water was poured in, determine how long the process has been going on.
   (*b*) If the graduate cylinder had *not* been empty and you cannot determine how much water was in it to begin with, can you make a statement as to the maximum or minimum duration of time this process has been going on?
2. Explain the technique used in determining the age of the solar system.
3. Describe the observations which indicate that the chondrules are older than the rest (matrix part) of the chondrite meteorites?
4. A rock is whirled on the end of a string 3 ft long. Its velocity is such that it makes one revolution in ½ second. The string is then shortened to 1 ft while the rock continues to whirl.
   (*a*) What is the speed of the rock when the string is 3 ft long?
   (*b*) According to the principle of conservation of angular momentum, what is the speed of the rock after the string has been shortened to 1 ft?
   (*c*) What is the period of revolution of the rock with the 1-ft string?

Chapter 9

1. What would be the horsepower rating of the sun if it were to:
   (*a*) double its diameter but maintain the same temperature?
   (*b*) double its temperature but maintain the same diameter?
   (*c*) double both its diameter and its temperature?

2. Explain by a diagram how the flash spectrum can be used to determine the heights of various elements in the solar atmosphere, as well as the heights of elements in various degrees of excitation.

3. Explain why a dark hydrogen filament appears bright when seen at the sun's limb projected against the dark sky.

4. List the similarities and differences between the photosphere, the chromosphere, and the corona.

5. Under what conditions can the chromosphere produce a bright-line spectrum?

6. How much energy is emitted per unit area each second by:
   (*a*) a toaster filament, $T = 1{,}000°K$
   (*b*) a furnace, $T = 2{,}000°K$
   (*c*) a star, $T = 4{,}000°K$
   (*d*) a hot star, $T = 16{,}000°K$

7. Find the wavelength (in Ångstroms) of the peak of the radiation curve for each of the four objects in question 6.

Chapter 10

1. Given the angle of parallax, find the distance of the following stars in parsecs:

(*a*) 61 Cygni	$p = 0.29$ seconds of arc
(*b*) Capella	$p = 0.07$ seconds of arc
(*c*) Tau Ceti	$p = 0.27$ seconds of arc

2. Find the distance in light years of each star in question 1.

3. A star has a proper motion of 0.5 seconds of arc per year. How many years will it be before the position of that star on the celestial sphere changes by 1°?

4. From the absolute magnitudes given, determine how much more luminous one star is than the other.

   (*a*) Beta Canis Majoris, M = −3.6
   Mira (Omicron Ceti), M = −2.6
   (*b*) Alpha Centauri, M = +4.1
   Delta Velorum, M = +0.1
   (*c*) Wolf 294, M = +10.9
   Castor (Alpha Geminorum), M = +0.9
   (*d*) Theta Aurigae, M = −0.2
   Beta Cassiopeiae, M = +1.8

5. Find the absolute magnitude of the following stars whose apparent magnitudes m, and distances d, are:
   (*a*) $m = 6.0$; $d = 5.0$ parsecs ($\log d = 0.70$)
   (*b*) $m = 6.0$; $d = 50.0$ parsecs ($\log d = 1.70$)
   (*c*) $m = 12.0$; $d = 5.0$ parsecs ($\log d = 0.70$)
   (*d*) $m = 12.0$; $d = 50.0$ parsecs ($\log d = 1.70$)

6. Two stars have the same luminosity but one has a brightness in the sky equal to nine times that of the other. If the brighter star is 25 parsecs from the sun, how far away is the other star?
7. Describe how the distance of a star can be estimated from a study of its spectrum.
8. Describe how the following stellar characteristics are related:
   (*a*) temperature, size, luminosity
   (*b*) luminosity, brightness in our sky, and distance from the sun
   (*c*) mass, luminosity, and temperature (for main-sequence stars)
   (*d*) mass and diameter (for main-sequence stars)
   (*e*) distance from the sun and proper motion
9. From this list of stars plot an H-R diagram.
   1. Sirius, A1, M = +1.5
   2. Alpha Crucis, B1, M = −2.7
   3. Ross 775, M4, M = +11.3
   4. Beta Pegasi, M2, M = −0.9
   5. 82 Eridani, G5, M = +5.3
   6. Sirius B, A5, M = +11.4
   7. Polaris, F8, M = −3.4
   8. Rigel, B8, M = −6.5
   9. Procyon, F5, M = +2.6
   10. Canopus, F0, M = −5
   11. Epsilon Indi, K5, M = +7.0
   12. Antares, M1, M = −4

Chapter 11

1. Draw a typical light curve and velocity curve for a Cepheid, the former directly above the latter. Now below the velocity curve draw a third curve to show generally how the diameter of the star changes.
2. What observations must be made before the distance of a Cepheid can be estimated?
3. How was it possible to change the period-brightness relationship for Cepheids in the Small Magellanic Cloud to a more general period-luminosity relationship?
4. Explain how the astronomer detects an expanding shell of hot gas about a star. How can he determine its velocity of expansion?

Chapter 12

1. For each of the following situations, indicate how a star's central temperature, central pressure, and diameter would change:
   (*a*) a decrease in the central pressure
   (*b*) an increase in the central temperature
   (*c*) an increase in the rate of energy generation
   (*d*) an increase in the rate of energy radiation from the surface
2. Explain why stars with larger mass form the upper part of the main sequence, and those with smaller mass form the lower part.
3. With the help of the H-R diagram explain the changes in temperature and luminosity of a star of moderate mass as it evolves from a parcel of gas and dust in a nebula.

4. What would be the differences in the brightest stars of a very young and a very old galaxy?
5. Cite observational evidence in support of the current theory of stellar evolution.
6. Calculate the pressure in atmospheres of one mole of a gas if that gas is confined to a volume of one liter (about one quart) at a temperature of $T = 300°\text{K}$ (about 80°F)?

Chapter 13

1. Why must the distance of a visual binary be known before the sum of the masses of the system can be determined?
2. What must be observed before the masses of the individual stars of a visual binary can be determined? Why must this observation be made?
3. (*a*) Capella is a visual binary system; the period of the system is 0.285 year and the separation of the two stars is 0.85 A.U. What is the sum of the masses of the two stars?
    (*b*) The mean orbital radius of the brighter member of the Capella system is 1.27 times that of the fainter member. What is the individual mass of each member?
4. The terms relative orbit, apparent orbit, and true orbit are used in the study of binary systems. What is the difference between each?
5. What would be the difference in the light curve of an eclipsing binary system if the major axis of its orbits were at right angles to the line of sight rather than parallel to it?
6. Why is the color index of an M-type star positive and that of a B-type star negative?
7. Discuss the principal differences between open clusters and globular clusters.

Chapter 14

1. What is the difference in spectra between an emission nebula and a reflection nebula? Why does this difference exist?
2. Is it possible over the next billion years or so for nebulae that are now dark to become bright? If so, what must happen?
3. Describe the process of fluorescence.
4. Explain why “forbidden” lines may be seen in emission nebulae.
5. Explain the process by which the 21-cm line is emitted. Why has this radio line achieved so much attention from the radio astronomer?
6. How can an astronomer detect the presence of clouds of neutral hydrogen; of ionized hydrogen?

Chapter 15

1. Explain the coordinate system that is most convenient for studies of our galaxy.
2. What observations have led to our understanding of the general shape of the Milky Way galaxy and to a knowledge of the location of the sun with respect to the center?
3. Which objects can best be used to trace the spiral arms of our galaxy? How can each of these objects best be observed?
4. Describe the general shape of the magnetic field in that spiral arm of our galaxy which contains our sun.
5. Discuss the "galactic fossils" of our Milky Way galaxy.

Chapter 16

1. List in order the methods of distance determination used by astronomers, beginning with the method that can be used only for the nearest stars and ending with the method used for the most distant galaxies.
2. Discuss three possible shapes our universe might take.
3. Outline the steady-state theory and the evolutionary theory of the universe. Compare the predictions made by each that lend themselves to observations, and that might be used to support one theory over the other.
4. If the number of quasars increases with increasing distance from our Milky Way galaxy, which of the two main theories of the universe would this observation support and why?

Index

Aberration, chromatic, 24–25
  spherical, 24
Absorption spectrum, 29–31, 54–55
Accretion, 192
Albedo, 64
  Earth, 120
  interstellar particles, 368
  Mars, 120
  Mercury, 64
  minor planets, 153
  moon, 64
  Venus, 120
Aldebaran, 6, 351–352
Andromeda galaxy, 281, 391–395
  Cepheids, 281
  distance, 281
  RR Lyrae stars, 281
Angle of parallax, 242–243
Angstrom unit, 15
Angular momentum, 190–191
  conservation, 190
  in galaxies, 424–425
  in solar system, 191–192
Antares, naming, 6
  position, 111
Aphelion, 98
Apogee, 98, 133
Apparent orbit, 334
Association, stellar, 357–360
Astroblemes, 166
Astronomical unit, 60
  relation to mile, 100
Atmosphere, Earth, 68–70
  giant stars, 343
  Jupiter, 79–80
  Mars, 73–74

Atmosphere, Mercury, 62–63
  moon, 120
  Saturn, 84
  Venus, 60–67
  white dwarfs, 323
Atomic mass, 46, 211
Atomic nucleus, 45
Atomic number, 46, 211
Atomic structure, 45–47, 49–52
Aurorae, 72, 224, 288
Autumnal equinox, 109

Balloon observations of sun, 213
  of Venus, 67
Balmer continuum, 370
Balmer series, 53–54
Barred spiral galaxies, 396–397
Bernard's star, 246, 349
Beta Canis Majoris stars, 317–318
  on H-R diagram, 279
Beta Lyrae, 345–348
Binary stars, 330–348
  eclipsing, 340–343
  evolution, 344–345
  orbits, 333–337
  spectroscopic, 337–340
  visual, 330–337
Bode's law, 149–150
Bonner Durchmusterung, 8
Brahe, Tycho, cometary observations, 168

Calendar, Gregorian, 6
  Julian, 6
Carbon cycle, 210–211, 305, 309
Carina-Cygnus spiral arm, 409
Cascading, 54
Cassini's division, 84
Celestial equator, 108–109
Celestial poles, 109, 111–112
Celestial sphere, 9, 109
Center of mass, binary stars, 331
  Earth-moon system, 117–118
Centripetal force, 143
Cepheids, 271–280, 284
  H-R diagram, 279
  light curve, 272, 273, 275
  number of, 284
  period-luminosity relation, 276–278
  Population I and II, 278
  pulsation, 273–276, 320
Cepheids, spectra, 274, 279
  velocity curve, 276
Ceres, 150
Chemical compounds, in comets, 172
  in interstellar material, 378
  in stars, 251
  in sun, 207
Chondrites, 161–164
  age, 197
Chondrules, 161–164
  age, 197
Chromatic aberration, 24–25
Chromosphere, *see* Sun
Clusters, globular, 360–363
    color-luminosity diagram, 318–320
    as distance indicators for galaxies, 436
    Hercules cluster, 361
    location in galaxy, 361, 406, 419, 420
    in Magellanic Clouds, 397–398
    number of, 361
    Omega Centauri, 360–361
    Population II, 362
    recent star formation, 362
  open, 350–357
    color-luminosity diagram, 353–355
    evolution, 354–360
    formation, 356–357
    Hyades, 351–352, 355
    identification of members, 351–353
    location in galaxy, 363, 392, 397, 407
    Pleiades, 354–357
    Population I, 362
    proper motion, 351–352
    Ursa Major, 351
Cluster variables, *see* RR Lyrae stars
Coalsack nebula, 374, 375
Color excess, 378, 409
Color index, 353–354, 378
Color-luminosity diagrams, 355
Comets, general, 168–180
    appearance, 170–172
    association with meteor showers, 174–176
    behavior near sun, 170–171
    collision with Earth, 166
    collision with moon, 124–125
    coma, 170, 172, 173, 175
    Earth passing through tail, 173

Comets, general, in history, 168–169
    Jupiter's family, 169
    mass, 174
    and micrometeorites, 175–176
    nucleus, 170, 172, 174–175
    orbits, 169
    periodic, 169
    spectra, 172
    tail, 170–172, 175, 178–179
  particular, 1843, 172
    1862 III, 174
    1866 I, 174
    Arend-Roland, 148, 177–178
    Giacobini-Zinner, 174
    Halley's, 169, 171, 172, 173–174, 176
    Holme's, 172
    Ikeya-Seki, 171, 179–180
    Morehouse, 172, 173
    Mrkos, 178–179
    Wirtanen, 174, 175
Conservation of angular momentum, 190–191
Constellations, 7–8
  see also star charts
Continuous spectrum, *see* Spectrum
Convective envelope, 212, 307, 308
Coordinate systems, celestial, 109–111
  galactic, 407–408
  terrestrial, 110
Core, convective, 209, 307, 311, 312–313
  Earth's, 68, 69, 158
  helium, 307, 311, 312, 322
  red giant, 312–313
  sun's, 211–212
Corona of sun, 137, 202, 232–236
  appearance during eclipse, 137, 232–235
  composition, 230–232
  spectrum, 208
  temperature, 235
Coronagraph, 227
Coronal rain, 229
Cosmogony, 429
Cosmological principle, 440, 456
Cosmology, 429
Crab nebula, 289–293
Craters, lunar, Alphonsus, 122, 125, 128–130
    Archimedes, 122, 123, 124
Craters, lunar, Aristillus, 122, 123, 124
    Autolycus, 122, 123, 124
    collapse, 127–128
    Copernicus, frontispiece, 122
    dark haloed, 128–130
    Flamsteed, 122, 126
    Julius Caesar, 122, 131
    Tycho, 132
  Martian, 76, 77
  terrestrial, Barringer, 165
    Chubb-New Quebec, 165
    Siberian, 166
Cygnus arm, 409

Dark globules, 184, 190, 314–315
Dark-line spectrum, *see* Absorption spectrum
Dark nebulae, *see* Nebulae
Day, apparent solar, 106, 107, 111
  sidereal, 106, 111
Declination, 110
Degenerate matter, 265
  cores of red giants, 312–313
  density, 265
  white dwarfs, 265, 321–324
Deimos, 140–141
Delta Cephei, 271–280
Density, 80
  interstellar material, 379
  Jovian planets, 189
  Orion nebula, 372
  sun's, 201–202
  supergiants, 265
  terrestrial planets, 189
  white dwarfs, 265
Deuterium, 210
Discrete radio sources, 381
Distance, measurement of, brightest stars in galaxy, 435–436
    Cepheids, 276, 378, 435
    globular clusters, 436
    heliocentric parallax, 104–105, 242–244, 434
    largest galaxies in clusters, 436
    RR Lyrae stars, 280–281, 435
    spectroscopic parallax, 258–262
    supernovae, 295, 436–437
Distance units, astronomical unit, 60, 100
  light year, 243
  parsec, 243

Diurnal (motion), 1, 7, 96
Doppler broadening, 235
Doppler effect, 31–33
  radar, 60–62
  relativistic, 458
Double stars, *see* Binary stars
Dust, intergalactic, 434
  interplanetary, 180
  interstellar, 379–380

Earth, 68–73
  age, 188
  albedo, 120
  atmosphere, 68–70
  axis of rotation, 107–108
  collision with a comet, 166
  craters, 165–166
  data, 90
  density, 90, 189
  distance from sun, 60
  internal structure, 68, 69
  magnetic field, 70–72
  magnetic storms, 225
  magnetic tail, 72–73
  oblateness, 102–103
  orbital velocity, 98–99, 106
  precession, 111–113
  proofs of revolution, 104–106
  proofs of rotation, 102–103, 112
  revolution, 104–109
  rotation, 102–104
  seasons, 107–108
  size, 90
  tides, 138–139
Earthlight, 120
Eastern elongation, 65
Eclipses, 135–138
  annular, 136
  frequency, 135
  in history, 137–138
  lunar, 135, 137
  partial, 135–136
  stellar, 340–343
  total solar, 116, 135–138, 234
Eclipsing binaries, 340–343
  annular, 340
  gaseous streams, 344–345, 347
  limb darkening, 341
  partial, 342
  total, 340
Ecliptic, 106
  and lunar nodes, 135

Electromagnetic spectrum, *see* Spectrum
Electron, 45–47
  capture, 370
  free, 308, 370
  temperature, 370
Electron spin, 45
Electron transitions, 50–56, 207–208
  and fluorescence, 369
  forbidden, 371
  free-free, 382
  and the 21-cm line, 383–384
Ellipse, 97–98, 144–145
Elliptical galaxies, *see* Galaxies
Emission nebulae, *see* Nebulae
Emission spectrum, *see* Spectrum
Energy, 48–49
  kinetic, 48
  potential, 48, 305
Energy generation in stars, 210–211
Energy levels in atoms, 49–50
Energy-mass relation, 211, 443
Energy transport, 306–308
Equator, celestial, 108, 407
  galactic, 407
  terrestrial, 108
Equinox, autumnal, 109
  vernal, 108
Escape velocity, 63, 90
  of Earth, 63
  of Mercury, 63
  of moon, 120
  of sun, 202
Evolution of galaxies, 423–425
  of Milky Way galaxy, 420–422
  of stars, evidence, 314–321
  of stars, theory, 306–308
  of universe, 455–456
Evolutionary track, 311
Expanding universe, 453–455, 457
Extragalactic systems, *see* Galaxies

Filament, dark, *see* Sun
Fireballs, 155
Flare stars, 381–382
Flares, solar, 222–224
Fluorescence, 369
Focal length, 23
Focal plane, 19, 20
Focus, Newtonian, 22
  prime, 21
Forbidden lines, 371

Foucault pendulum, 102
Frame of reference, 96
Fraunhofer lines, 27–28
Free-free transitions, 382

Galactic coordinate system, 407–408
Galactic equator, 407
Galaxies, angular momentum, 423–425
  barred, 396–397
  clusters, 430–434
  distances, 434–438
  distribution, 429–434
  dwarf, 399–400
  elliptical, 388, 390, 391–393
  evolution, 423–425
  exploding, 400–403
  irregular, 390, 397–398
  mass, 423–425
  red-shift, 438–440
  spiral, 390, 392, 393–396
Galaxy, *see* Milky Way galaxy
Galilean satellites, 141
Galileo, 9, 18, 121, 141
Gas laws, 304–305
  applied to stars, 305–306
Gegenschein, 180
Geocentric, 97
Geodesic, 446
Giant stars, 257
Globular clusters, *see* Clusters, globular
Graben, 130–132
Granules on solar surface, 211–214
Gravitational red-shift, 460
Gravity, law of, 101
Greenhouse effect, 66
Gregorian calendar, 6

h and λ Persei, 328, 354, 355
H-R diagram, 257–258
Halley's comet, *see* Comets, particular
Heliocentric, 97
Helium flash, 313
Herbig-Haro objects, 302, 314, 316, 357
High velocity stars, 247, 417
Hyades, 351–352
  color-luminosity, 355
Hydrogen, atomic model, 46–47
  spectrum, 29, 53–55
  21-cm line, 383–384
Hyperbola, 144–145

Image, 19–20
Inclination of orbit, 334
Inferior conjunction, 65
Interferometer, beam, 248
  radio, 39, 40
Intergalactic gas, 434
Interplanetary dust, 180
Interstellar dust, 379–380
  location in galaxy, 380, 419
  scattering of star's light, 378
Interstellar gas, 377–380
  absorption lines, 346, 378
  density, 379
  detection by radio telescopes, 378, 382–385
  location in galaxy, 378, 418–420
Interstellar lines, 345–346, 377–379
Inverse square law, 259
Ionization, 51–52, 54, 208, 382
Isotope, 46

Jovian planets, 181
  density, 189
Julian calendar, 6
Jupiter, 79–84
  atmosphere, 79–80, 82
  bands, 80, 81
  data, 90
  density, 80–82, 90, 189
  distance from sun, 79, 90, 99
  Great Red Spot, 80, 81
  internal structure, 82
  mass, 80, 90
  oblateness, 79, 82
  radio studies, 82–84
  revolution, 79
  rotation, 79
  satellites, 16–17, 141
  size, 79, 80
  temperature, 79

Kant-Laplace theory, 191–192
Kepler's laws, 97–100, 143
  as applied to binary stars, 331–332
Kepler's nova, 284, 292
Kuiper's theory, 193–196

Laplace, *see* Kant-Laplace theory
Latitude, galactic, 407–408
  terrestrial, 110
Leap year, 5–6
Lens, 18, 24

Light, polarized, 15–16
  reflection, 18
  refraction, 18
  velocity, 16–18
  wavelength, 14
Light curve, of Delta Cephei, 272–275
  of eclipsing binaries, 340, 342
  of novae, 285
  of U Geminorum stars, 299
Light year, 243
Limb, 61
Limb darkening, in sun, 225
  in stars, 341
Local Group of galaxies, 432
Longitude, galactic, 407–408
  terrestrial, 110
Long-period variables, 281–283, 284
  placement on H-R diagram, 279
Luminosity, of stars, 254, 256–257
    related to mass, 264
    related to spectra, 262
Luminosity function, 266–267
Lunar eclipses, 135, 137
Lunar nodes, 135
Lyman continuum, 370
Lyman series, 53

Magellanic Clouds, 397–398
  composition, 397–398
  satellites of Milky Way galaxy, 398
Magnitudes, absolute, 255–257
  apparent, 33, 254–255
  photographic, 353
  sun's, 256
  visual, 353
Main-sequence stars, 257
  age of, 316–317
  spectra compared with giants, 261
Maria, lunar, 123–128
  Martian, 76
Mars, 73–79
  albedo, 120
  atmosphere, 73–74, 78
  canals, 75
  clouds, 74, 75, 78–79
  color, 76, 78
  craters, 76, 77
  data, 90
  density, 90, 189
  distance from sun, 73, 90
  magnetic field, lack of, 74
  maria, 76
Mars, Mariner IV studies of, 73–74, 76, 77
  mountains, 74
  polar regions, 74–75
  revolution, 73, 90
  rotation, 73, 90
  satellites, 140–141
  seasonal changes, 74, 75
  size, 73, 90
  temperature, 73
  wave of darkening, 74–75
Mass, 100
Mass, center of, 118, 331
Mass-energy relation, 211, 443
Mass-luminosity relation, 263–265
  applied to stars and planets, 349
Mercury, 60–64
  albedo, 64
  atmosphere, 62–63
  data, 90
  density, 90, 189
  distance from sun, 60, 90
  escape velocity, 63
  radar studies of, 60–62
  revolution, 60
  rotation, 60–62
  size, 60, 90
  temperature, 64
Meridian, prime, 110
  terrestrial, 110
Metastable states, 371
Meteorites, 49, 154–168
  achondrites, 161
  age, 158, 188
  chondrites, 161–164, 197
  chondrules, 161–164, 197
  composition, 157–158, 164
  crystals, 158–159, 164
  relation to minor planets, 158–159
Meteoritic craters, *see* Craters
Meteors, 154–157
  brightness, 155
  frequency during night, 155, 157, 166
  height, 156
  number, 155, 167
  radar detection, 156–157
  size, 155
  sporadic, 166
  velocity, 156
Meteor showers, 166–168
  radiant, 167
  relation to comets, 174–176

Michelson-Morley experiment, 440–442
Micrometeorites, 174–176
Milky Way galaxy, 9, 403–417
  age, 324–325
  evolution, 420–425
  location of sun, 405–407, 409
  magnetic field, 418–420
  nucleus, 407, 418
  rotation, 409–417
  size, 407
  spiral arms, 409, 410, 415, 417–420
Minor planets, 150–154
  albedo, 153
  discovery, 150, 151–152
  distribution, 152
  origin, 154, 195–196
  relation with meteorites, 158–159
  size, 153
Mira, 283
Molecular bands, in comets, 172
  in interstellar gases, 378
  in long-period variables, 283
Month, sidereal, 133–134
  synodic, 133–134
Moon, albedo, 64
  atmosphere, 120–121
  changes in temperature, 132
  collapse craters, 127–128
  collision with comet, 124–125
  dark-haloed craters, 128–130
  distance from Earth, 117
  Earthlight, 120
  eclipses, 135–138
  erosion, 127, 130
  far side, 132
  graben, 130–132
  lava flows, 123, 125
  libration, 133–134
  map, 122
  maria, 123
  motion, 133–135
  occultations, 120
  origin of craters, 122
  phases, 118–120
  radar studies, 132
  rays, 123
  rills, 128
  secondary craters, 123–125
  surface material, 126–127
  surface rocks, 126
  temperature, 132
Nebulae, 368–377
  birth of stars, 357–360
  bright, 368–372
  dark, 372–377
  density, 372
  diffuse, 368
  emission, 368–372
  planetary, 296–298
  reflection, 368
  spectra, 368, 370
Nebulae, particular, Coalsack, 374, 375
    Horsehead, 372, 373
    Lagoon, 184
    Loop, 294
    M-16, 366
    near Rho Ophiuchi, 373, 374
    Rosette, 315
    see also Crab nebula and Orion nebula
Nebular hypotheses, 190–196
Negatively curved space, 452
Neptune, data, 90
  density, 90, 185
  discovery of, 87
  distance from sun, 88, 90
  revolution, 88
  rotation, 88
  satellites, 142
  size, 88, 90
Neutron, 45, 46
Newton, Sir Isaac, 4, 21, 100
Newtonian focus, 22
Newton's laws, 100–101, 142
Nodes, lunar, *see* Lunar nodes
Normal spiral galaxy, 392–396
North celestial pole, 6
North star, 6, 111–112
Novae, 284–288
  dwarf, 298–299, 343–345
  light curve, 285
  luminosity, 285
  spectra, 286–288
Nuclear reactions, 210–211
Nucleus, of atoms, 45–46
  of comets, 170, 175
  of galaxies, 393, 418
  of Milky Way galaxy, 405, 407, 418

Objective lens, 18
Oblateness, cause, 102–103
  of Earth, 102–103
  of Jupiter, 82

Occultation, lunar, 120
Open clusters, *see* Clusters, open
Orbits, of comets, 169–170
  of eclipsing binaries, 342–343
  foci, 98
  of minor planets, 151, 152
  of planets, 97–99
  of satellites, 135, 142–146
  of spectroscopic binaries, 337–340
  of visual binaries, 331–333
Orion nebula, 26–27, 357–360
  density, 372
  radio observations, 357
  star association, 357–360

P Cygni stars, 295–296
Parabola, 24, 144–145
Parallax, angle, 242–243
  heliocentric, 104–105, 242–243
  mean, 416
  spectroscopic, 258–262
Parsec, 243
Peculiar velocity, 247
Penumbra, of shadow, 135–136
  in sunspot, 216
Perigee, 98, 133
Perihelion, 98
Period, 16
Periodic comets, *see* Comets
Period-luminosity relation, 276–278
Phases, of the moon, 118–120
  of Venus, 9, 65
Phobos, 140
Photocell, 38
Photographic magnitude, *see* Magnitudes
Photoionization, 54, 64
Photometry, 33–36
Photons, 52
Photosphere, *see* Sun
Plages, 200, 222, 223
Planck's law, 206
Planetary nebulae, 296–298
Pleiades, 354–357
Pluto, 88–90
  data, 90
  discovery, 88
  distance from sun, 88
  former satellite of Neptune, 88–90
  revolution, 88
  rotation, 88
  size, 88
Pluto, temperature, 88
Population I objects, ages, 422
  blue giants, 281
  Cepheids, 278
  location in galaxies, 393, 420–422
  open clusters, 362
Population II objects, ages, 410
  Cepheids, 278
  elliptical galaxies, 393
  globular clusters, 362
  location in galaxies, 281, 297, 393, 420–422
  RR Lyrae stars, 280
Positively curved space, 452
Precession, 6, 111–113
Pressure broadening, 56, 260–261
Prime focus, 21
Prominences, 226–233
  arch, 228, 229, 230–231
  coronal rain, 229
  loop, 229, 233
  quiescent, 229, 233
  surge, 229, 232
Proper motion, 244–245
Proton, 45–47
Proton-proton reaction, 210, 305, 309
Protoplanet, 193
Protostar, 308
Proxima Centauri, 243
Pulsating stars, *see* Beta Canis Majoris Stars; Cepheids; Long-period variables; RR Lyrae stars; and Semiregular variables

Quasars, 458–461

RR Lyrae stars, 280–281
  distance indicators, 280–281, 406, 435
  in elliptical galaxies, 393
  in globular clusters, 360
  on H-R diagram, 279
  as high velocity stars, 417
  location in Milky Way galaxy, 406, 420, 422
  number of, 284
  period, 280
  spectra, 280
  in stellar evolution, 320
Radar reflected, from Mercury, 60–62
  from meteors, 156–157
  from Venus, 64, 100

Radial velocity, 33
  annular variation in, 106
  of galaxies, 438–440
  of quasars, 458
  of stars, 33, 244, 245
Radiant of meteor shower, 167
Radiation laws, 202–206
  Planck, 203, 206
  Stefan-Boltzmann, 203–205, 279
  Wien, 205–206
Radiation pressure, in nebular disc about protosun, 196
  in stars, 308, 321
Radiative zone, 212, 307
Radioactive half-life, 187–188
Radioactivity, 187–188
Radio map of sky, 381
Radio radiation, from Crab nebula, 290, 292
  from galaxies, 400–403, 404
  from ionized hydrogen, 382
  from Jupiter, 82–84
  from neutral hydrogen, 383–384
  from Orion association, 357
  from quasars, 460
  from spiral arms, 412–414, 415
  from sun, 224
Radio telescopes, *see* Telescopes
Radius vector, 98
Recurrent novae, 298
Red-dwarf stars, 257, 324
Red-shift, of galaxies, 438–440
  gravitational, 460
Reflecting telescopes, *see* Telescopes
Reflection nebulae, 368
Refracting telescopes, *see* Telescopes
Refraction of light, 18
  atmospheric, 70
  simple lens, 24
Relativity, general, 444–445
  special, 440–444
Resolving power of telescope, 37–40
Retrograde motions, 106
  of planets, 64, 106
  rotation, 64, 66
  of satellites, 142
Revolution, of Earth, 104–109
  of moon, 133
  of planets, 60
Right ascension, 111
Roche's density, 194
Roche's limit, 194
Rocket astronomy, 294–295
Rotation, of Andromeda galaxy, 393–396
  of artificial satellites, 154
  of Earth, 102–104
  of Jupiter, 79
  of Mercury, 60
  of Milky Way galaxy, 400–407
  of minor planets, 154
  of moon, 133
  of the planets, 60
  of protoplanet, 194, 195
  of sun, 216–218
  of Uranus, 86–87
  of Venus, 69, 90

Sagittarius arm, 409, 410
Satellites, artificial, 142–146
  of the Earth, *see* Moon
  of Jupiter, 16–17, 141
  of Mars, 140–141
  of Saturn, 141
  of Uranus, 142
Saturn, 84–86
  atmosphere, 84
  data, 90
  density, 84, 90, 189
  distance from sun, 84, 90
  oblateness, 84
  revolution, 84
  rotation, 84
  satellites, 141
  size, 84, 90
  temperature, 84
Saturn's rings, aspects, 85
  composition, 84–85
  gaps in, 84
  rotation, 84–85
  size, 84
  spectrum, 84, 85
  transparency, 85
Scattering, in Earth's atmosphere, 70, 137
  in interstellar gases, 378
  in nebulae, 368
Schmidt telescope, 25
Seasons, of Earth, 107–108
  of Mars, 74, 75
Semiregular variables, 283–284
Shell stars, 295–296
Sidereal day, 106, 111
Sidereal period of moon, 133–134

Sidereal time, 106, 111
Sirius, to ancient Egyptians, 5
  distance, 256
  magnitudes, 256
  member of Ursa Major cluster, 351
  orbit, 336
  spectral type, 335
  white-dwarf companion, 335–336
61 Cygni, 255, 348, 349
Solar eclipse, 116, 135–138, 234
Solar flares, 222–224
Solar nebular disc, 191, 192, 193
Solar radiation curve, 204
Solar spectrum, 206–209
Solar system, age, 186–188
  definition, 59
  origin of, 188–198
Solar time, 106, 111
Solar wind, 236
  and comets, 170–171, 236
  and Earth, 72–73
  and moon, 127, 130
Solstice, summer and winter, 109
Space, Euclidean, 446
  negative curvature, 452
  non-Euclidean, 446–453
  positive curvature, 452
  zero curvature, 452
Spectral energy curve, 203–206
Spectral line, 28, 52
  intensity, 207
Spectral-luminosity diagram, *see* H-R diagram
Spectral sequence, 249–253
Spectral-temperature relation, 250–252
Spectrograph, 26–28
Spectroheliogram, 221
Spectroscopic binary, 337–340
  gaseous streams, 344, 347
  masses, 338–339
  velocity curves, 337–338
Spectroscopic parallax, 258–262
Spectrum, absorption (dark-line), 29, 54–55
  comparison, 29–30
  continuous, 27–28, 56
  electromagnetic, 35–36
  emission (bright-line), 28–29, 30, 52–54
  flash, 209
  formation of, 27, 50–56
Spicules, 226, 227
Spiral galaxies, barred, 396–397
  composition, 393
  normal, 393–396
  rotation, 394–396
  spiral-arm structure, 417–420
Star clouds, 240
Star clusters, *see* Clusters, open; Clusters, globular
Star trails, 2, 7
Stars, absolute magnitudes, 255–257
  apparent magnitudes, 33, 254–255
  associations, 357–360
  brightness, 254
  chemical composition, 267, 425
  color excess, 378, 409
  color index, 353–354, 378
  distances, 242–244
  energy generation, 210–211
  formation of, 190, 197, 308–309, 314, 348–349
  giants, 257
  interiors, 209–212, 306–309
  luminosities, 253–257
  main-sequence, 257
  mass, 262–263, 265, 331–333
  naming, 6–8
  nearest, 243
  sizes, 247–249, 343
  supergiants, 257, 265
  temperatures, 249–253
  white dwarfs, 257, 265
Stefan-Boltzmann law, 203–205, 261
Stellar associations, 357–360
Stellar mass, determination of, 331–333
Stellar parallax, *see* Parallax, heliocentric
Stellar velocities, of high velocity stars, 247, 417
  peculiar, 247
  radial, 33, 244–245
  relative, 244–245
  tangential, 244–245
Summer solstice, 109
Sun, absolute magnitude, 256
  angular momentum, 191–192, 196
  apparent magnitude, 256
  apparent motion, 5
  central pressure, 212
  central temperature, 212
  chromosphere, 137, 202, 226–232
  composition, 206
  corona, 137, 202, 232–236

Sun, dark filaments, 200, 222, 223
  density, 201
  effect on Earth, 107–108, 224, 236
  energy output, 202–203
  escape velocity, 202
  flares, 222–224
  flash spectrum, 209
  Fraunhofer lines, 27–28
  granulation, 211–214
  interior, 209–212
  limb darkening, 225
  magnitudes, 256
  mass, 201
  motion in galaxy, 416–417
  photosphere, 202
  plages, 200, 222, 223
  prominences, 137, 226–233
  radio emission, 224
  rotation, 216–218
  size, 201
  spicules, 226, 227
  surface temperature, 205–206
Sun's motive relative to Milky Way galaxy, 410–411, 416–417
Sunspots, 214–221
  cycle, 218–221, 236
  groups, 214–216
  magnetic field, 218–221
  relation to prominence activity, 229
  solar rotation, 216
  temperature, 216, 221
Supergiant stars, 257
Superior conjunction, 65
Supernovae, 288–295
  distance indicators, 295, 436
  luminosity, 288, 295
  remnant, 289–295
Synchrotron radiation, 290, 402
Synodic period of moon, 133

T Tauri stars, 314–316, 357
Tangential velocity, 244, 245
Tektites, 159, 160
Telescopes, radio, 36–41
  reflecting, 21–24
  refracting, 18–21
  Schmidt, 25–26
Temperature, absolute, 205 fn
  electron, 370
Terminator of moon, 121
Terrestrial planets, 181
  density, 190
Tidal forces, on Earth, 138–139
  on protoplanet, 194–195
Tides, of moon, 138–139
  neap, 139
  spring, 139
  of sun, 139
Time, sidereal, 111
  solar, 111
Titan, 141
Tree rings, 236
Triton, 142
True orbit, 334
Turbulent broadening, 214
21-Cm line, distance indicator, 412–414
  origin, 383–384
Tycho's nova, 284, 292

U Geminorum stars, 298
Umbra, in shadow, 135–136
  in sunspot, 216
Universe, evolutionary theory, 455–456
  expansion, 439, 453–455
  origin, 455–457
  pulsating, 456
  shape, 445–453
  steady-state theory, 456–457
Uranus, 86–87
  data, 90
  density, 90, 189
  discovery, 86
  distance from sun, 86, 90
  revolution, 86
  rotation, 86–87
  satellites, 142
  size, 86
  temperature, 86
Ursa Major cluster, 351

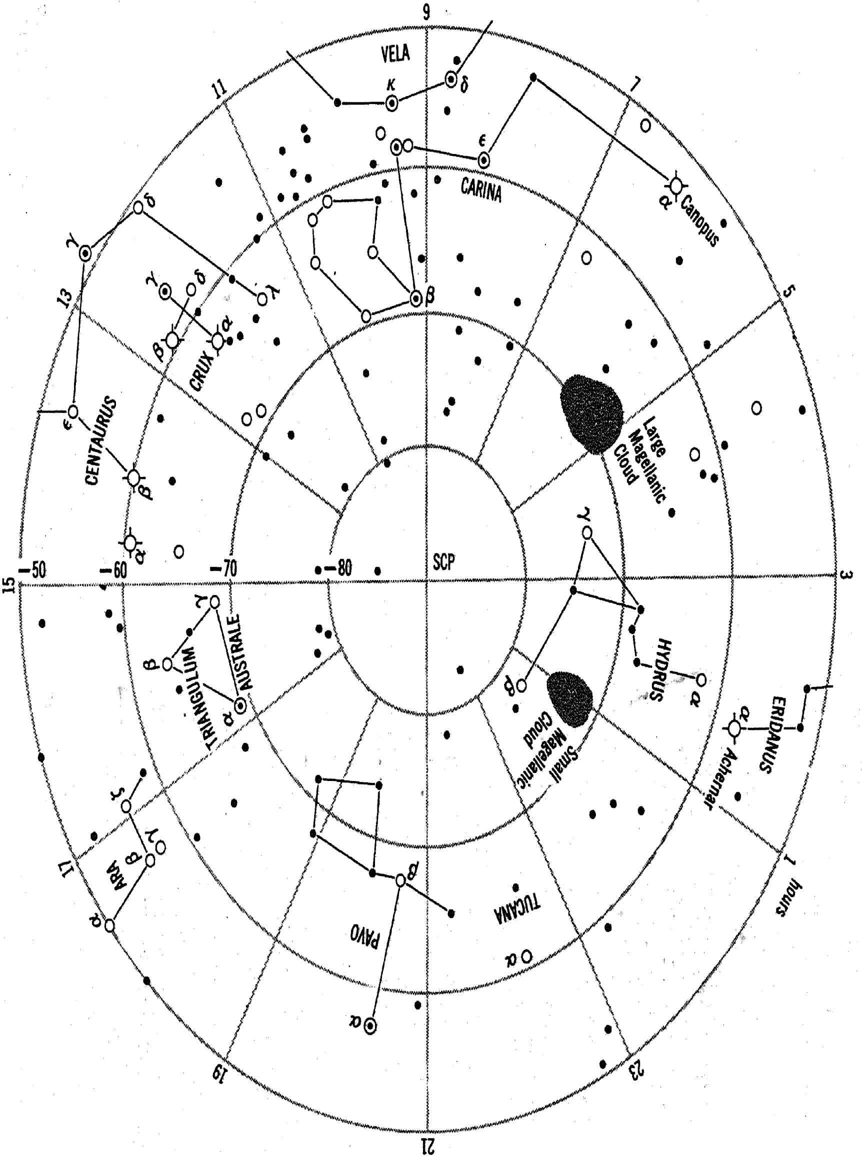
Van Allen radiation belts, 70, 71
Variable stars, Beta Canis Majoris, 317–318
  Cepheids, 271–280
  on H-R diagram, 279
  Mira, 283
  RR Lyrae, 280–281
  semiregular, 283
Variation in length of day, 104
Vega, naming, 6
  as north star, 112–113
  position, 111
  size, 248

Velocity, escape, *see* Escape velocity
Velocity curves, Cepheids, 274, 275
  long-period variables, 283
  spectroscopic binaries, 337–339
Velocity-distance relation for galaxies, 434–438, 439
Velocity of light, 16–17, 440–444
Venus, 64–68
  albedo, 120
  atmosphere, 66–67
  data, 90
  density, 90, 189
  distance from sun, 64, 90
  Mariner II studies of, 67
  radar studies of, 64, 66
  revolution, 66, 90
  rotation, 66, 90
  size, 64, 90
  temperature, 66, 67
Vernal equinox, 108
Visual binaries, *see* Binary stars
von Weizsäcker, theory, 192–193

Wandering poles, 104
Wavelength, 14
Western elongation, 65
White dwarfs, 257, 265, 321–326
  atmospheres, 322
  degenerate gas, 265, 321–323
  density, 265
  gravitational forces, 322
  mass-diameter relation, 322
  observational evidence, 323–324
  spectra, 323
  temperature, 323–325
Wien's law, 205–206, 282
Winter solstice, 109
Wolf-Rayet stars, 296
Work, 47–48

Zeeman effect, 218
Zenith, 70
Zero-age main sequence, 309
Zero curvature space, 452
Zodiacal light, 180

THE SOUTHERN CONSTELLATIONS



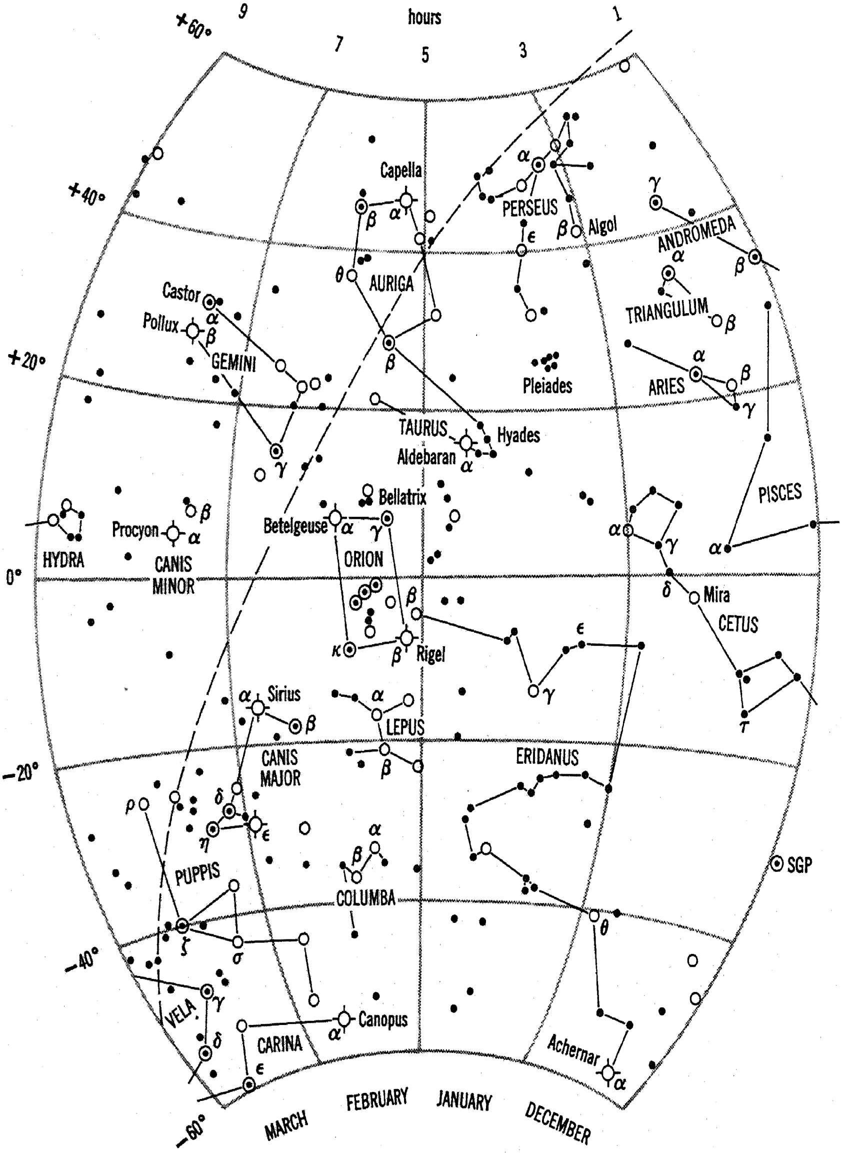
-✧- brighter than magnitude 1.5

◉ between magnitudes 1.5 and 2.5

○ between magnitudes 2.5 and 3.5

● fainter than magnitude 3.5

THE WINTER CONSTELLATIONS



⌖ brighter than magnitude 1.5

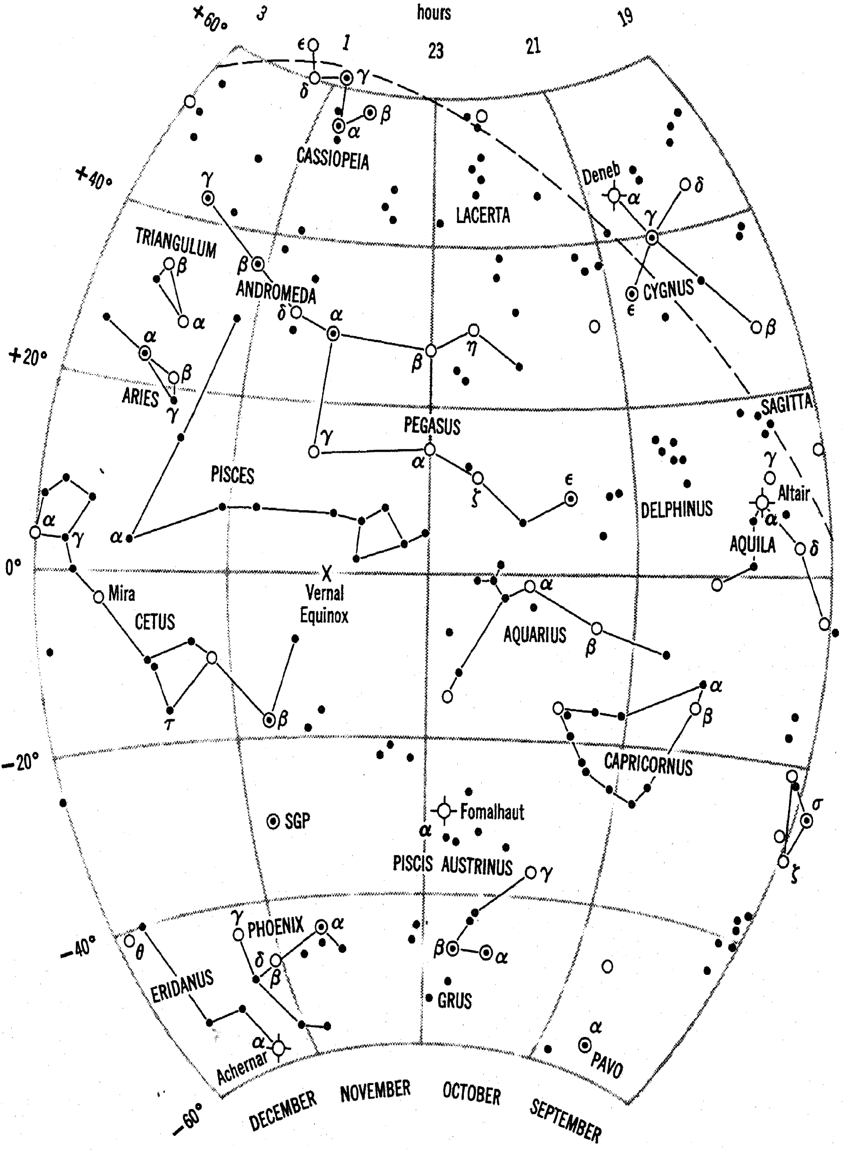
⊙ between magnitudes 1.5 and 2.5

○ between magnitudes 2.5 and 3.5

● fainter than magnitude 3.5

– – – plane of Milky Way Galaxy

THE AUTUMN CONSTELLATIONS



⌖ brighter than magnitude 1.5

◉ between magnitudes 1.5 and 2.5

○ between magnitudes 2.5 and 3.5

● fainter than magnitude 3.5

––– plane of Milky Way Galaxy

UNIVERSAL LIBRARY

102 835

UNIVERSAL LIBRARY